Public Knowledge and Perceptions of Large-Scale Assessments

A Thesis

Submitted to the Faculty of Education

In Partial Fulfillment of the Requirements

for the Degree of

Master of Education

University of Prince Edward Island

Yan Yan

Charlottetown, PE

January, 2016

ABSTRACT

Large-scale assessment has been used as an effective tool for government organizations to justify the well-being of educational systems in terms of accountability, gatekeeping, instructional diagnosis, and monitoring student achievement. The purpose of this study was to examine public knowledge and perceptions about large-scale assessments and thereby explore the accountability function of large-scale assessments. An online questionnaire combined with a paper and pencil questionnaire was distributed to residents in a small Canadian province using a nonprobability purposive sampling technique combined with convenience sampling. A total of 515 questionnaires were completed. The overall findings revealed that public was knowledgeable about students' most recent performance on the Programme for International Student Assessment but not the most recent performance on the Pan-Canadian Assessment Program. The public's perceptions towards large-scale assessments were in the middle of the scale and there was no statistically significant differences based on parental status, educational attainment, or cultural affiliations.

Key words: large-scale assessment, common assessment, accountability, stakeholders, Prince Edward Island

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Tess Miller, who encouraged me and gave me strength in this academic journey, for her expertise, direction, support, patience and academic rigour to guide me through my thesis. Dr. Tess Miller has an ever lasting influence on my experience as a master student for all her suggestions, criticism, and most importantly—the enthusiasm in helping me persevere in persuing this thesis.

I would also like to thank my committee member, Cindy Wood, for her valuable advice, expertise in Prince Edward Island's large-scale assessment program, and in the entire process of completing my thesis. To the examining committee, I would like to express my sincere appreciation for your time and patience reviewing the thesis.

Finally, I would like to express my thanks to my family and all my friends who have supported me through this journey. Thank you!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAC	Alberta Assessment Consortium
ADEP	Alberta Diploma Examination Program
ANOVA	Analysis of Variance
ATP	The Achievement Testing Program
CMEC	The Council of Ministers of Education, Canada
EFTO	Elementary Teachers' Federation of Ontario
EQAO	Educational Quality and Assessment Office
OECD	The Organization for Economic Co-operation and Development
ICT	Information communication technology
LSA	Large-scale Assessment
NCTM	National Council of Teachers of Mathematics
OISE	The Ontario Institute for Studies in Education
OSSLT	Secondary School Literacy Assessment
PCAP	The Pan-Canadian Assessment Program
PE	Prince Edward Island
PEITF	Prince Edward Island Teachers' Federation
PIAAC	The Programme for the International Assessment of Adult Competencies
PISA	The Programme for International Student Achievement
PS-TRE	Problem solving in technology-rich environments
SAIP	SchoolAchievement Indicators Program
SAT	Scholastic Assessment Test
SEF	School Effectiveness Framework

CHAPTER ONE

The challenges to individuals and societies imposed by globalization and modernization are widely acknowledged and apparent (van Davier, Gonzales, Kirsch, & Yamamoto, 2013; Murry, 2003; Schleicher, 2010). Increasing diversity and interconnected populations are intertwined with rapid technological advancements, which has placed demands on societies to balance economic growth with the sustainability of natural environments, individual prosperity, and a reduction of societal inequalities. To meet these demands, the Organization for Economic Co-operation and Development (OECD) has advocated for the development of a population's knowledge, skills, and competencies through education (OECD, 2014a). The OECD has also emphasized the need for education systems to equip citizens with the skills necessary to achieve their full potential and participate in an increasingly interconnected global economy. Ideally, this would be a catalyst for better employment opportunities and, subsequently, a better quality of living, which is a central preoccupation of policy makers around the world (OECD, 2014a).

Reports from the OECD's survey of adult skills, known as the Programme for the International Assessment of Adult Competencies (PIAAC), emphasized the importance that training and skills had on salary so much that "highly skilled adults are twice as likely to be employed and almost three times more likely to earn an above-median salary than poorly skilled adults" (OECD, 2013a, p.12). Conversely, and although obvious, inadequately trained adults are restricted from applying for better-paying and more rewarding occupations, thereby placing demands on social assistance as they are more likely to be unemployed. In the twenty-first century, the skills and qualifications of all citizens are pivotal to national and

regional economic development.

In our accelerating and globalized world, economic success is no longer measured against national standards alone, but rather against the best performing and most rapidly improving economic and culturally advanced societies. These societies are situated in healthy education systems that prepare students to enter a knowledge economy calling for transferable and versatile skills. Contemporary calls for advancement in the way schools prepare students for the future cite the need for critical thinking skills given the uncertainty of the future workforce (Schleicher, 2010). This uncertainty is attributed to the emergence of problems that cannot be foreseen. To be prepared for the demands of the globalized and interconnected economy, students of the future must embrace the basic skills taught in mathematics, science, and language arts and then use them to synthesize, interpret, analyze, and to think critically. Recently, educators have also advocated that knowledge and skills related to information communication technology (ICT) should be included in the kindergarten to 12 education system to provide students with the tools to adapt to their future careers and lifestyle—especially in digital reading and problem solving in technological environments (Davier, Gonzalez, Kirsch, & Yamamoto, 2013). Given that education plays a vital role in the prosperity of national and international growth, such a colossal responsibility has not been left in the hands of teachers, school principals, or school boards. Rather, the OECD along with national, and provincial governing bodies (e.g., Ministries of Education) have each implemented large-scale assessments (LSAs) as a means of providing information for public accountability. This accountability strategy provides information about student achievement such as identifying students and schools in need. As Klinger (2005) noted, without LSAs, "students in low achieving areas may not be exposed to the

same important learning goals as other students, making their subsequent competitiveness weaker for work or getting accepted into college or university." Miller (2015) stated that responsible governance of education uses LSAs to gauge how well they educate their children and thereby provides a proxy measure of the growth, prosperity, and sustainability of Canadian jurisdictions. Based on the outcome of LSAs, many countries have reformed educational policies and revised national and/or regional strategic planning objectives with the aim of enhancing their intellectual capital—the collective achievement of their elementary and secondary students by drawing from LSA reports (Volante, 2008).

There have been a number of studies exploring both teachers' and students' accountability towards learning (Koch, 2011; Miller, 2013; van Barneveld, 2008). However, little is known about the broader public's perceptions about the accountability purpose of LSAs or other LSA purposes such as gatekeeping or instructional development, and the role LSA plays in the economic development of a community. This is an important component in understanding LSAs because an absence of awareness or little regard towards LSAs suggests the threads weaving the nuances of the LSA framework together may be weak or flawed. In such a context, LSAs may dwindle down to nothing more than an excessively hefty government expense that bears nothing more than denigration from the public.

Purpose

The purpose of this study was to examine public perceptions of LSAs in terms of knowledge about LSAs and the perceptions they hold for these instruments. The research questions posed in this study were:

1. What are the public's understandings of students' performance on large-scale assessments?

2. What are the public's perceptions towards the utility of large-scale assessments as instrument of accountability and how do perceptions differ based on participants' parental status, educational attainment, and cultural affiliation?

If the public has little awareness of student performance on LSAs and holds little value towards LSAs, then LSAs are likely not being used as a means to keep educators accountable, let alone to pressure policy makers or educators to make use of any other possible functions of LSA data. Given the relatively low-stakes nature of many LSAs, combined with negative disposition towards LSAs espoused by teachers' unions (Wright, 2013), it was hypothesized that the public are not well informed about LSAs and, subsequently, consider LSAs to be of little value. Furthermore, it is posited that the accountability purpose of LSAs is influenced by the stake of the LSA where the higher the stake, the greater the accountability.

Significance of the Study

Examining public perceptions and comparing them with the OECD's stated purposes of LSAs acts as an indicator of whether LSAs are viable instruments of accountability and, subsequently, good proxy measures of a community's growth, prosperity, and sustainability. Furthermore, there is little empirical research on perceptions about LSAs among the Canadian public, and this study aims to contribute to this body of knowledge. Findings will be useful to all educational stakeholders: Economic leaders (e.g., Chamber of Commerce), education policy makers, teachers, and administrators.

Background of Large-scale Assessment

The province of Prince Edward Island (PE) was selected for study primarily because student achievement in this province has persistently lagged behind other provinces and territories on national and international LSAs (CMEC, 2001, [PISA: reading, mathematics, science]; 2004, [PISA: mathematics, science, reading]; 2007 [PISA: science, mathematics, reading]; 2008 [PCAP: reading, mathematics, science]; 2010 [PISA: reading, mathematics, science]; 2011 [PCAP: mathematics, reading]; 2013 [PISA: mathematics, reading, science]). In addition, the relatively low-stakes nature associated with PE's LSAs may be influencing a culture that has little value for student achievement on LSAs. It is timely to conduct a study on public perceptions of LSAs before the implementation of provincial high-stakes LSAs slated to be introduced in the 2015–2016 school year. This section provides a backdrop of the different levels of LSAs and contrasts some of the high and low achieving provinces. A detailed examination of PE student achievement on LSAs is presented in a separate chapter dedicated to the context of LSA in PE.

LSAs at the International Level

As previously noted, the Organization for Economic Co-operation and Development (OECD) administers two international assessments. The first is the Programme for International Student Assessment (PISA), which began in 2000 and assesses the achievement of Grade 10 students in three key domains of reading, mathematics, and science prior to the completion of their compulsory schooling (OECD/UNESCO-UIS, 2003). The OECD also administers the Programme for the International Assessment of Adult Competencies (PIAAC), which began in 2011. PE participates in both of these assessments and their achievement is discussed below.

PE achievement on the PISA. PE students have consistently scored at the bottom of the PISA scale in five cycles of the assessments compared to students from other provincial jurisdictions. For example, in the science domain, PE ranked near the bottom of the 10 Canadian provinces[1] over the five cycles of administration (see Table 1.1). The same is true for PE student achievement in mathematics and literacy (see Tables 1.2 and Table 1.3).

Table 1.1

Summary of PISA Scores in Science

	2000	2003	2006	2009	2012
Top jurisdictions	Alberta	Alberta	Alberta	Alberta	Ontario
Average scores	546	529	550	545	527
PE rankings (out of 10)	9th	10th	9th	10th	10th
Average scores	508	489	509	495	490

Note. Adapted from CMEC PISA reports 2001, 2004, 2007, 2010, and 2013.

Table 1.2

Summary of PISA Scores in Mathematics

	2000	2003*	2006	2009	2012*
Top jurisdictions	Quebec	Alberta	Quebec	Quebec	Quebec
Average scores	550	549	540	543	530
PE rankings (out of 10)	8th	10th	10th	10th	10th
Average scores	512	480	501	487	485

Note. Adapted from CMEC PISA reports 2001, 2004, 2007, 2010, and 2013.
* In 2003 and 2012, the main domain of the PISA was mathematics. The average score and the ranking were based on student performance on combined mathematics in 2003. In 2012, the average score and the ranking were based on student performance on composite mathematics.

[1] The three Canadian territories do not participate in PISA.

Table 1.3

Summary of PISA Scores in Reading

	2000	2003	2006	2009*	2012*
Top jurisdictions	Alberta	Alberta	Alberta	Alberta	British Columbia
Average scores	550	543	535	533	542
PE rankings (out of 10)	8th	10th	10th	10th	10th
Average scores	517	495	497	486	490

Note. Adapted from CMEC PISA reports 2001, 2004, 2007, 2010, and 2013.
* In 2009, the average score and the ranking of reading were based on student performance on combined reading. In 2012, the average score and the ranking were based on student performance on composite reading.

PE achievement on PIAAC. The survey of adult skills was first introduced by the OECD as part of the Programme for the International Assessment of Adult Competencies (PIAAC) in 2011. It mainly assesses 16 to 65 year olds in literacy, numeracy, and problem solving in technology-rich environments (PS-TRE). In addition, cognitive, social, learning, and physical skills are also evaluated and complimented with a background questionnaire (OECD, 2013). In Canada, ten provinces and three territories took part in the survey. The results revealed that Alberta and Ontario were above the OECD average in literacy, while Prince Edward Island was at the OECD average. In terms of numeracy, all 13 Canadian provinces and territories were performing at or below OECD's average. PE ranked third after Alberta and British Columbia, which were at the OECD's average. With respect to PS-TRE, except for Nunavut and Newfoundland and Labrador, all provinces and territories were above or at the OECD's average. PE was at the average of the OECD, ranking ninth among the 13 jurisdictions.

LSAs at the National Level

At the national level, the Council of Ministers of Education, Canada (CMEC) also acknowledged the need for Canadians to know how well their education system is meeting

the needs of students and those of society (CMEC, 2015). This governing body implemented the Student Achievement Indicators Program (SAIP) between 1996 and 2004 to measure the ability of 13 year olds (Grade 8) and 16 year olds (Grade 10) in the areas of science, mathematics, reading, and writing (CMEC, 1999). The SAIP was replaced by the Pan-Canadian Assessment Program (PCAP) in 2007, which now focuses only on the assessment of Grade 8 (Secondary II in Quebec) students in three key domains: science, mathematics, and reading (CMEC, 2014). This national measure provides provincial Ministers of Education with a basis for examining their own provincial/territorial curriculum and other aspects of their school systems. More importantly, the outcome of LSAs is to create a high quality public education system that provides a foundation for professional and advanced skill development to build a strong and competitive workforce ensuring sustainable economic development and future prosperity (Cirtwill, Clifton, & D'Orsay, 2002).

Since school programs and curricula vary from jurisdiction to jurisdiction across Canada, some argue that comparing results from these programs is too complex and can not provide feedback on students' progress to improve classroom instruction in a timely manner (Falk & Ort, 1998; Newton, Tunison, & Viczko, 2010; Stiggins, 2001; The Forum of Educational Accountability, 2007; Volante, 2006). Other researchers recognized that LSAs are a worthy task in terms of weighing the advantages and shortcomings of students' learning outcomes across the country (CMEC, 2014; Taylor & Tubianosa, 2001).

When comparing provinces, Alberta led its western counterparts, scoring at or near the top on the PCAP science (i.e., PCAP, 2007, 2010, 2013). They also scored high on the PCAP of reading (i.e., second in 2010; third in 2007 and 2013) and mathematics (i.e., third in 2007, 2010, and 2013) (CMEC, 2014). Alberta has two assessment programs that are

administrated annually: the Achievement Testing Program (ATP) and Alberta Diploma Examination Program (ADEP, Alberta Education, 2015a). The ADEP is administered to students enrolled in Grade 12 for all courses (i.e., English Language Arts, French Language Arts, Biology, Chemistry, Mathematics, Physics, Science, and Social Studies). The diploma examination mark is worth 30% of the total mark, compared with 70% of the school mark (Alberta Education, 2015a). Likewise, the province of Quebec, which edged above the western provinces on the last PISA (2012) and also stood among the top scoring jurisdictions in all domains on the previous two cycles of the PISA, determined students' final grades in a similar manner (Ministère de l'Éducation, du Loisir et du Sport, 2014). By contrast, the Atlantic provinces do not have such high-stakes LSAs and coincidently or not, they have repeatedly scored poorly on national LSAs.

PE students, in particular, have consistently demonstrated difficulty on national assessment. In the 2007 PCAP, when reading was the main domain, PE students performed significantly lower than other Canadian students with a mean score of 460 compared with the Canadian mean score of 500 (CMEC, 2008). Although there was an improvement in reading, as can be seen in the 2013 PCAP, PE students' achievement was still lower than the Canadian mean score (CMEC, 2014a). In mathematics, PE students not only had lower mean scores compared with other jurisdictions, but they also had a large proportion of students (13%) in the lowest level (Level 1) in comparison to the Canadian average of 9% for students scoring in the lowest level. The same is true for the upper end of the scale where PE had a small percentage (29%) of students scoring in Levels 3 and 4 compared with the Canadian average of 47% in these upper two levels on the 2010 PCAP (CMEC, 2011). This situation did not change PE's ranking until the most recent assessment (i.e., PCAP 2013),

when PE showed considerable improvement scoring an average of 492 compared with Canada's overall mean of 507. In science, there was a similar pattern of performance in that PE had consistently ranked lower than other Canadian students (CMEC, 2008, 2014a, 2014b).

Provincial/Territorial LSAs. From a Canadian perspective, education is a provincial/territorial responsibility. Therefore, each province also creates its own curriculum and corresponding LSA program. Provincial/territorial assessments most often occur in Grades 3, 6, and 9 and are typically regarded as low-stakes instruments since the results have no influence on teachers' pay or promotions, nor do they have any influence on students' promotion at key stages in the education system (Klinger, Maggie, & D'Angiulli, 2011; Klinger, DeLuca, & Miller, 2008; Volante, 2014). Higher-stakes assessments are found in the senior grades (i.e., Grade 10 literacy assessment in Ontario and PE, Grade 12 exit tests in Alberta), and the scores on these LSAs are either used for promotion (e.g., Grade 10 Literacy Assessment in Ontario and PE [forthcoming]) or as a component of students' final grades (e.g., exit tests in Alberta, Grade 11 mathematics assessment in PE [forthcoming]) (Alberta Education, 2015a; Prince Edward Island Department of Education, Early Learning and Culture, 2015a).

PE implemented provincial LSAs in 2007, making it the last Canadian province to do so. PE has not finished initiating all levels of assessments, nor achieved the systematic or transparent approach seen in Ontario and Alberta (note: this topic will be expanded on the contextual chapter) (PE, Department of Education, Early Learning and Culture, 2015b). Similar to other jurisdictions in Canada, PE introduced LSAs at the key stages of learning in Grade 3 (reading, writing, mathematics), Grade 6 (reading, writing, mathematics), and Grade

9 (reading, writing, mathematics). Educators and theorists typically consider PE's provincial LSAs as low-stakes, since, with the exception of the Grade 9 literacy and mathematics LSA, none of the LSA scores are used in determining students' final grades (Miller, 2013). Even in the case of the Grade 9 assessment of mathematics, in which a 10% stake was supposed to be applied (i.e., 10% of students' LSA score counts towards students' final score), there has been little regulation or oversight guiding this policy. Therefore, teachers and schools vary in their practice of calculating final scores in Grade 9 mathematics (Miller, 2013).

PE students' LSA performance in mathematics reflected the achievement on national and international LSAs. From 2010 to 2013, the percentage of students who met the provincial requirement ranged from 68% to 65% in Grade 3. When examining the qualified students in Grade 6, the proportion decreased to 56% (2013 provincial assessment result). For Grade 9, there had been an increase in the average score from 59 to 66 between 2008 to 2012. In terms of reading, the percentage of students who met the standard[2] ranged from 62% to 72% from 2007 to 2013 through Grade 3 to Grade 9.

It is important to note that public data released in the annual reports by the PE Department of Education and presented on their website are different from the data obtained through email correspondence with the Instructional Development & Achievement Office of the PE Department of Education. The data available on the website report the percentage of students who met or approached the standard, which was a composite of the number of students who met the standard and a proportion of students whose performance approached

[2] Note: exemplars of the standards/expectations or other such criteria are not made available to the public.

the standard. In comparison, the reports from the Instructional Development Office present the actual proportion of students who met the standard and the percentage of students who approached the standard. Furthermore, proportions of students who experienced difficulty meeting the standard were also presented in these exclusive reports. The percentage of students who are approaching the expected standard ranged from 3% up to 11%, which is discussed in chapter three focusing in greater detail on the context of PE.

Another issue that needs to be pointed out is that provincial assessment data did not include students who had special learning needs and who were eligible to use adaptations in their tests. Unlike the Educational Quality and Assessment Office (EQAO) in Ontario and the Alberta Assessment Consortium (AAC), which provides provincial data to guide improvement in school leadership and updated publications about using assessment to help students and parents achieve the standards including the assessment of special needs students, PE does not have such a rich resource for the public to access.

High-stakes assessments will not occur in PE until the implementation of an electronic Grade 10 literacy assessment (paralleling the Ontario literacy test administered by the EQAO)—referred to as the Secondary Literacy Assessment in PE—and a Grade 11 Mathematics assessment (i.e., Secondary Mathematics Assessment), which are due to be fully implemented in 2019 and 2016 respectively. These two LSAs will bring significant change to the use of LSAs in PE since success on the Grade 10 literacy assessment will become a graduation component, and the Grade11 mathematics assessment is slated to count for 25% of students' final score in mathematics. Since these are significant factors affecting both students' passage from secondary to post-secondary institutions (or the world of work), these LSAs will be viewed as higher stakes than the LSAs in Grade 3, 6, and 9. Further,

they are likely to influence people's perceptions of LSAs given that the gatekeeping function of the Grade 10 literacy assessment and the 25% associated with the LSA of Grade 11 mathematics is likely to raise students' and parents' awareness of the importance of this assessment. In general, the higher the stake, the more attention people will pay to the task. This is true in teaching as much as it is in any other profession or aspect of one's life (Driesler, 2001; Phelps, 1998).

Given the absence of substantial stakes associated with the existing LSAs, perceptions of LSA practices in the province of PE may not be overly positive such that people may harbour negative dispositions towards LSAs as influenced by the media based on the LSA position of the local teachers' union (Horne, 2008). It is possible that these perceptions may indirectly influence student performance on LSAs by creating a culture that is shaped by a misunderstanding of the purposes and significance that LSAs play in ensuring the growth, prosperity, and sustainability of a province. This is a problem not only for the wider public but also for schools and educators, since the extent to which the accountability framework of LSAs (as espoused by Klinger, DeLuca, & Miller, 2008) is able to hold those responsible for education accountable is likely to also be influenced by the stakes associated with the instrument. While it is evident from the new LSAs being implemented that policy makers in PE believe there is a need for higher-stakes LSAs, it is nevertheless, hypothesized that this is not a view yet shared by the public, and similar to the opposition encountered when other provinces introduced high-stakes LSAs (Burger & Krueger, 2003), it is likely that PE will encounter some opposition to LSAs as they are implemented in the upcoming year.

Conceptual Framework

The four functions of large-scale assessments (LSAs) serve as the conceptual framework for this thesis. Drawing on Nagy's work in 2000, the first three functions are described as the ability to hold the education system accountable with the task of educating children, ensuring those who are prepared move forward, and providing individual teachers with feedback on the effectiveness of classroom instruction relative to standards articulated on the assessments. Klinger, DeLuca and Miller extended LSA's fourth functionality of monitoring student progress (DeLuca, Klinger, and Miller, 2008). These four functions have been abridged to accountability, gatekeeping, instruction diagnosis, and monitoring student progress.

The earliest role that LSAs played was for admission or graduation, which was considered a *gatekeeping* function. An example of this would be the American Scholastic Aptitude Test (Nagy, 2000). A more recent example of the gatekeeping function would be Ontario's Secondary School Literacy Assessment (OSSLT), which students are required to pass to receive their Grade 12 diploma (EQAO, 2015). To date, PE does not have any high-stakes LSAs hence the gatekeeping function of LSAs does not apply.

LSAs have also been used to evaluate whether schools were proficient in preparing qualified students who are capable of achieving successful careers and contributing to society. It is because of this role that post-secondary institutions document the employment pathways of their graduates. This role of LSAs has morphed into what we know as the *accountability* function. Accountability might be better understood as system-wide accountability (Earl, 2010), wherein all stakeholders in education are responsible for various aspects of educating children.

The latter two roles LSAs play are to provide feedback for instructional diagnosis and monitoring student progress. Ideally, all educational stakeholders would use statistical reports based on LSA data to review curriculum delivery and make modifications as needed to improve the quality of teaching and learning. More specifically, teachers and school administrators would particularly benefit from reviewing LSA results and using their findings to refine programs, teaching methods, and strategies. Research focusing on issues associated with the four functions of LSA is presented in the following chapter that provides a synthesis of the literature in this area.

CHAPTER TWO

Literature Review

Literature focusing on large-scale assessments (LSAs) is organized into the four functions of LSAs: accountability, gatekeeping, instructional diagnosis, and monitoring student achievement. Following an examination of literature related to LSA functions, literature focusing on perceptions is synthesized. Perceptions held by four key stakeholders in education included teachers, principals, parents, students, and the general public.

Accountability

Educational accountability is a method of governance ensuring that educational stakeholders maintain the quality of the schools and instruction within it. Accountability is not only focused on students and teachers, but also extends to policy makers, administrators, and parents (Linn, 2000, 2003; Volante, 2007). Ultimately, educational accountability encompasses students' ability to meet their future responsibilities in society, which requires the effort of all stakeholders. This section highlights the means by which LSAs can be used to ensure accountability and discusses related problems associated with accountability in the wider Canadian context as well as in PE.

Research has shown that LSA scores are effective levers in influencing students' accountability when a portion of students' LSA scores are used to determine their overall grade in a course (Fushell, 2011). Some Canadian provinces have developed practices that call for a portion of the LSA score to be used in calculating a student's final score in a course. Since education in Canada is a provincial responsibility, there appears to be no national structure or format to this practice. In Newfoundland, for example, 20% of students' grades are based on their Grade 9 mathematic assessment score (Fushell, 2011).

In Ontario, where polices do not mandate the use of a particular weighting of LSA scores, studies have shown that some teachers do not use LSA scores at all, whereas others use up to 20% of the score to determine students' final grades (Koch, 2011). In Alberta, the diploma examination program, which is a compulsory provincial assessment for high school graduation and accounted for 50% of students' final score, with the other 50% allocated by the teacher (Alberta Education, 2015). This weighting recently dropped to 30% in the 2015–2016 school year. In PE, teachers are required to integrate 10% of students' score from their LSA of Grade 9 mathematics to determine students' final standing in the course (Miller, 2013). As previously mentioned, the practice varied from teacher to teacher due to the absence of policies and procedures to guide and monitor the practice.

Although the practice of using LSA scores to determine students' final standing in a course varies from one province to the next, LSAs hold students accountable for their own learning. The rationale behind this practice was related to student motivation. If students were motivated to engage in LSAs, it was believed that they would prepare for, and complete the assessment to the best of their ability, thus providing a more accurate reflection of their true ability in the domain being measured (van Barneveld, Pharand, Ruberto, & Haggarty, 2013). This form of external motivation has raised concerns for theorists in this area who noted that students did not appreciate this assessment-oriented motivation. Instead, students were pushed to memorize knowledge and spent too much time on drilling activities due to the increased emphasis on the high-stake nature of LSAs (Brophy, 2004; Lipnevich & Smith, 2008; Toshalis & Nakkula, 2012). Barneveld et al. (2013) argued a need to internally motivate students by building a sense of confidence in knowledge and a positive attitude towards a successful assessment (p. 52). Motivation appears to be the primary catalyst

behind incorporating students' LSA scores as part of their overall standing in a course. In van Barneveld and Brinson's study (2011) that focused on student perceptions in Grade 9 applied mathematics, 70% of the 15,831 students indicated that the practice of using LSA scores when calculating their final grade motivated them to take the test seriously. Disregarding the argument related to appropriateness of the practice, if and when LSAs are incorporated into student grades, it is important to communicate to all stakeholders, especially students, the purpose and value of LSAs along with teachers' expectations, how scores influence students' overall standing in a course, and the consequences of assessment outcomes (if any).

Although Canada has a shorter history with LSAs than the United States, and the extent to which those responsible for educating children are held accountable differs between countries, Nagy (2000) claimed that Canada's provincial governments have been putting increasing emphasis on the outcomes and the quality of the educational system similar to their American counterparts. However, LSAs in Canada have evolved considerably since Nagy's pivotal analysis of LSAs in 2000. For example, in 2000, Ontario had just commenced their LSA program, while PE had not yet implemented any provincial LSAs. Subsequently, the extent to which provincial governments, including PE, are attempting to improve the education system is largely unknown as a result of sparse literature examining accountability from the perspective of students, teachers, school boards, and the greater public.

Based on a study in Ontario, van Barneveld, Pharand, Ruberto, and Haggarty (2013) found that high LSA achievers were intrinsically motivated to exceed and receive personal satisfaction at working hard to score well. For low-achieving students, a lack of self-esteem

and low expectations for success decreased their motivation to succeed. These researchers concluded that if students understood the purpose and consequences of LSAs, they might be more inclined to exert more effort. These factors affecting students' accountability with LSAs appear to be underpinned by theories of motivation, which warrants a separate investigation into how it impacts students' accountability.

In another, more recent study examining teachers' LSA practices in PE, Miller (2013) discovered that although school boards had a policy for incorporating a percentage of students' LSA scores to determine their final scores in mathematics, the absence of supervision in guiding teachers' practices resulted in differences in how teachers interpreted the guideline. In addition, the absence of guidance in administering LSAs contributed to teachers' naivety towards LSAs. This likely influenced the accountability factor of LSAs in that teachers or schools were not being held accountable for adhering to the guidelines established by the province and that resulted in ad hoc practices (Miller, 2013).

Studies examining school boards' understanding of the extent that LSA scores are used to influence students' overall standing in a course were not found. The same was true for literature related to the public's understanding of accountability and LSAs. The emergence of literature in these areas is needed to determine the extent the accountability framework is working with these educational stakeholders.

It is important to note that incorporating LSA scores towards students' grades is not a required practice needed to extend the accountability framework to students and parents. Talking to parents about the importance of LSAs and sharing student and school scores with parents is another way of extending the accountability framework. However, little is known about the culture of LSAs within the classroom since it is difficult to conduct research within

school settings. Strategically, this study aimed to examine the accountability factor through surveying perceptions of the PE community, including the perceptions of parents with children in schools who might be able to provide insight into how accountability is conceptualized in a school or community setting.

There have been some concerns raised with respect to the accountability framework of LSAs. Teachers and some educational researchers (e.g., Nagy, 2000; Volante, 2004) are opposed to using LSAs to measure the well-being of the school system because of varying perceptions about the reliability[3] and validity[4] of the instrument, particularly given the ease by which the security of the tests can be compromised during administration (Simon, Ercikan, Rousseau, 2013; Fey & Smith, 2000; Wollack & Fremer, 2013). These concerns compromise the usefulness of LSA results with respect to the accountability function, given that some may view LSA results as inaccurate representations of students' learning.

This skepticism surrounding the accuracy of LSAs is heightened further when considering who is assessed and what assistance they receive, if any, because it is not well communicated (Koretz, 2003; Morison, McLaughlin, & McDonnell, 1997). Guidelines for accommodations and modifications do specify the allocation of extra time or use of scribes (Alberta Education, 2015b; PE Department of Education, Early Learning and Culture, 2015c; Education Quality and Accountability Office (EQAO), 2015). Other accommodations such as large-print versions of LSAs and different colour print are to be offered to students in need of visual aids (Alberta Education, 2015b). Similar accommodations are offered on national

[3] Reliability: Reliability is the degree to which an assessment tool produces stable and consistent results.

[4] Validity: In educational testing, validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests.

assessments (CMEC, 2010). Researchers have questioned the appropriateness of these accommodations and modifications for special needs students. Extended time was regarded as one of the common adaptations offered to students with special needs that presented some controversy. Runyan (1991) showed that extra time benefited the students with special needs by improving their performance on the LSA. However, Zuriff (2000) noted that both students with or without special needs could score better on LSAs under this accommodation, which was aligned with the findings from other studies (Elloitt & Marquart, 2004; Fuchs, Fushs, Eaton, Hamlett, Binkley, & Crouch, 2000; Lewandowski, Lovett, Parolin, Gordon, & Codding, 2007). Conversely, other studies showed no significant difference in the scores of students who had disabilities and students who were in non-inclusive classrooms (Demeris, Childs, & Jordan, 2007; Huber, Rosenfeld, & Fiorello, 2001; Sharpe, York, & Knight, 1994).

Other issues influencing the validity of LSAs were raised with the assessment of second language learners (Bolt & Thurlow, 2004; Elliott & Marquart, 2004; Fuchs, Fushs, Eaton, Hamlett, Binkley, & Crouch, 2000). It is likely there is a lot of misinformation about the amount of accommodations that these students receive, if any, and how their scores influence students' overall average for a jurisdiction. Koretz and Barton (2003) argued this point noting that the assessment of special needs students was plagued with issues and inconsistencies. In Ontario, for example, the Education Quality and Accountability Office (EQAO) included the provincial assessment results of students with special needs in the Ontario Secondary School Literacy Test (OSSLT) (EQAO, 2015). In PE, the decision to have second language learners write the Grade 3 and Grade 6 assessments in reading and writing is at the discretion of the teacher and school principal, which is made without providing documentation to the Department of Education (PE Department of Education,

Early Learning and Culture, 2015c). Given that the decision to withhold these students from writing the LSA is at the school level, it is likely that the general public is unaware of the practice. Additionally, in schools with a large number of immigrant families who are second language learners, low LSA scores are likely attributed to the poor scores from second language learners and, thus, the LSA is discounted since some may believe that it does not reflect the abilities of most students.

In summary, the discrepancies surrounding who is assessed and whether LSA scores from special needs students are included in the analysis of the general student population appears to cast a shadow on the accuracy of LSAs. Subsequently, this appears to influence the accountability of LSAs given that those who are responsible for educating children may feel absolved of any wrong doing if they view LSA scores as invalid measures.

In addition to these concerns, some researchers argued that meaningful school reform will only occur if educators have the autonomy and opportunities to use LSA data to identify students' needs and explore what teaching practices and methods will help meet these needs (Elmore, 2004 ; McTighe & Thomas 2003). Other researches have proposed various models of using LSA data to promote school and district improvement efforts (EQAO, 2014; Parke, 2012; Volante, 2007). In these calls for reform, school teams firstly identify learning goals, and then collect, analyze, and summarize data from multiple sources for the purpose of determining the extent of student learning. Next, they consider the causes of present achievement levels to design systemic actions that address root causes, promote continuous and enduring learning, and raise student achievement (Klinger, Maggie, & D'Angiulli, 2011). The accountability framework calls on educators within schools and related authorities to engage in this reform. Accountable educators would effectively use LSA data to identify

priorities, track results, and determine future actions to improve student learning. This, however, requires training and, in practice, is not always as effective as the theoretical model would suggest. Hence, this may also be a factor affecting the accountability framework in that little regard is held for LSA data because teachers may not feel empowered to use it due to an absence of knowledge or time to review and apply LSA findings to their teaching.

Accountability models in Canada generally follow this professional accountability framework, which was historically derived from global movements of school effectiveness (Normand, 2008). Nevertheless, there are challenges in this accountability framework, mainly due to the absence of guidance in interpreting the data, identifying systematic learning needs, and using the information to improve classroom instruction for all students. All of these actions require time and training for teachers, which, ironically, takes away from instructional time in the classroom, and means that even teachers who wish to use the LSA data are not always able to and, thus, the accountability function is once again compromised.

Some provinces have initiated provincial assessment programs to mitigate these problems. These programs are designed to monitor student educational outcomes, identify areas for improvement, and enact initiatives and practices that will result in increased student achievement (Klinger, DeLuca, & Miller, 2008; Ontario Ministry of Education, 2010). In Ontario, the School Effectiveness Framework (SEF) assumes that educators will monitor their own instructional effectiveness by engaging in analyzing their student assessment data obtained either from provincial results or school- or board-based common assessments (Ontario Ministry of Education, 2013). This belief is based on research that suggests one of the factors that seems to be associated with job satisfaction and sense of self-efficacy among teachers is the success of students (Leithwood, 2006; Moore & Esselman, 1994; Muijs &

Reynolds, 2002; Ross, 1992, 1998). The SEF document provided a relatively common process through which educators measured the effectiveness of their initiatives and, therefore, enabled them to take responsibility for improving the educational system from within, with its focus on increasing instructional effectiveness and data-based decision making through the identification of students' strengths and weaknesses. Ontario has a clear framework for how to use LSA data to enhance school performances (Ontario Ministry of Education, 2013), and, therefore, the tests are able to fulfill their accountability function for all stakeholders. Other provinces, such as Alberta, follow similar practice on school improvement (Government of Alberta, 2015c). However, PE falls short in this regard, with only one school act establishing the principles of education in the province, but there were no separate mechanisms of accountability in the public education system (PE Department of Education, Early Learning and Culture, 2015d).

Without LSAs, schools would never know where their students were in the scale of learning achievement and what problems prevented them from achieving acceptable performance in relation to students in other schools and across Canada. While recognizing that LSAs are key components that allow all stakeholders to monitor and guide student learning, there are many factors that compromise the validity of the instruments and thereby may absolve some stakeholders from being accountable.

Gatekeeping

The concept of gatekeeping refers to the means by which students who score well on certain examinations (usually LSAs) receive recognition and passage into other programs, while students who do not score as well are either kept back at the "gate" to repeat the program, re-test, or are directed into programs more suitable to their skill and knowledge

level. This section discusses the history of the gatekeeping function of LSAs and then offers a review of Canadian provincial jurisdictions' use of LSAs in a gatekeeping function. Lastly, a discussion about the differences between high- and low-stake LSAs and the potential problems associated with this function is presented.

The gatekeeping function of LSAs can be traced back to the 1930s and 1940s when governments required skilled workers but realized how uneven their skills were due to differences in the quality of education and qualifications of teachers (Nagy, 2000). LSAs were consequently developed by governments to ensure quality control of educational systems as well as the enhancement of teacher education programs to generate well-prepared teachers for the classroom. The education system then implemented LSAs as a means to admit qualified students to higher programs and certify students for the appropriate career pathways. This early method of being able to choose the most qualified students for advanced training became known as the gatekeeping function in education system. In the United States, the most common test that served as a gatekeeper was the Scholastic Assessment Test (SAT). Top achievers on the SAT typically earn entrance into the most prestigious colleges and universities. In Canada, the gatekeeping function of LSAs is evident in the practice of using LSA scores to determine students' overall standing in a course. Table 2.1 presents the percentage and type of LSA scores that are part of high school graduation requirements. The data have been drawn from government files from each jurisdiction in Canada.

Table 2.1
Types of LSA and Percentage of LSA Score Integration in Requirements of High School Graduation in Canada

LSA involved in jurisdictions	LSA score integration (%)
British Columbia	
Grade 10 Language Arts	20
Grade 10 Mathematics	20
Grade 10 Science	20
Grade 11 Social Studies	20
Grade 12 Social Studies	40
Grade 12 Language Arts	40
Alberta	
Grade 12 Diploma examination	30
Saskatchewan	
Departmental exams *	40
Ontario	
The Ontario Secondary School Literacy Test (OSSLT)	Compulsory assessment
Quebec	
Secondary V Language of Instruction (Grade 11)	Compulsory assessment
Secondary IV Mathematics (Grade 10)	Compulsory assessment
Secondary IV Science and Technology/Applied (Grade 10)	Compulsory assessment
Secondary IV History and Citizenship (Grade 10)	Compulsory assessment
New Brunswick	
Grade 9 English Language Proficiency Assessment	Compulsory assessment
Nova Scotia	
Grade 10 English	20
Grade 10 Mathematics/ Mathematics at work	20
Prince Edward Island	
Grade 9 Mathematics	10
Grade 10 Literacy	Compulsory assessment
Grade 11 Mathematics	25
Newfoundland and Labrador	
Grade 9 English Language Arts	20
Grade 9 Mathematics	20
Newfoundland and Labrador	
Grade 9 English Language Arts	20
Grade 9 Mathematics	20

Note. Data adapted from Department of Education of provincial governments.
* Department exams only administrated to students instructed by non-accredited teachers

Based on the information presented in Table 2.1, Ontario, Quebec, New Brunswick, and PE have at least one compulsory LSA that students need to take before graduation. According to different policies in each jurisdiction, students may repeat the exam as necessary to pass. Other provinces do not have a compulsory LSA, but do make use of LSA scores as a portion of students' overall standing in a course, with the exception of Saskatchewan, which does not include any LSA scores to determine students' standing in courses. Alberta, as noted previously, decreased the percentage stake of their final LSAs, from 50% to 30%. The reason for the decrease was to put more emphasis on classroom work and the school-awarded mark, since teachers work closely with students on a daily basis over a longer period of time rather than relying so heavily on exams lasting only a few hours (Alberta Education, 2015a).

As Klinger et al. (2008) noted, using LSAs to determine students' final standing in a course serves as another example of the gatekeeping function, as it is assumed that the higher the percentage of LSA score used to determine students' final standing in a course, the more impact that the LSA would have on stakeholders like students, parents, teachers, principals, policymakers, and, subsequently, the greater public. Students, for example, may not prepare for or respond to LSA questions to the best of their abilities if LSA scores do not influence their standing in a course, being promoted from one grade to the next, or graduating from secondary school (van Barneveld & Brinson, 2011). Further, it is unknown how much the LSA needs to be worth to influence student engagement or how much weight is needed before the LSA is considered a high-stakes assessment. It is interesting to note that the data showed that most LSAs involved in determining final grades were considered to be low-stakes but the criteria for determining the stake of the LSA is not well articulated.

Hence, the range of weights associated with LSAs in different provinces suggests that perceptions of accountability may vary from one jurisdiction to the other.

With respect to the present study, it is important to note that prior to the new provincial assessment of Grade 11 mathematics, scheduled for implementation in 2016, PE did not have any LSA that could function as a gatekeeper for senior high schools (i.e., grades 10 to 12) (PE Department of Education, Early Learning and Culture, 2015e). This could possibly be a major contributor influencing the public's absence of awareness about LSA purposes and functions.

Moving from the level of provincial policy to current practice in PE, there has been some discrepancy in the extent in which teachers and principals use LSA results to guide the development of school programs and instruction. In PE, evidence of teachers' skepticism towards LSAs was found in a study examining how Grade 9 mathematics teachers determined students' overall standing in a course (Miller, 2013). In this study, teacher practices varied considerably when examining a practice called for by the Department of Education to allocate 10% of students' performance on Grade 9 mathematics to determine students' final standing in a course. Individual teachers' perceptions towards the utility of the instrument tended to influence the degree to which they engaged in this practice.

This ambivalence can also be seen in other jurisdictions. In Newfoundland and Labrador, students write the Grade 9 English Language Arts and Mathematics provincial assessment. The marks of these assessments are worth 20% of students' standing in a course. Fushell's (2011) study indicated that teachers were not actually using the provincial test scores for determining students' final score. Instead, teachers were using questions based on their own criteria, using their own rating guides, and applying their own standard.

Considerable differences in the comparison of teacher scores and provincial scores were found. For the English Language Arts, the greatest discrepancy was when student scores were in the 50% to 60% range. This practice is similar to what Miller (2013) reported about a guideline used by PE's Department of Education that prevented students from failing a course as a result of a poor LSA score. It appears that teachers in Newfoundland and Labrador are following the same practice but without an official guideline. This inconsistency of using provincial assessment scores to determine students' final standing in a course may invalidate the gatekeeping function as well as the accountability function of LSAs given that students do not seem to be held at the gate due to poor performance on an LSA.

Another situation that raises concerns about the gatekeeping function is related to Ontario's Grade 10 literacy test. To receive a Grade 12 diploma, students are required to pass the Grade 10 test. As expected, a number of students in any given year are anticipated to be unsuccessful with this test. Students who perform poorly can write the literacy test a second time in the next academic term. In the event students are not successful the second time around, they can take a literacy course. Upon completion of the literacy course, students receive an equivalent standing to passing the literacy test (Ontario Ministry of Education, 2003). Upon first inspection, this LSA appears to have a gatekeeping function similar to other LSAs. However, the repeated attempts to write the LSA and the literacy course seem to negate any function related to gatekeeping. Research documenting the impact of LSAs with higher weights in the senior grades is needed to collaborate the feasibility of the gatekeeping function in public schools given that no research in this area was found.

Given the varied practice of integrating scores from LSAs to determine students' final standing in a course, the gatekeeping function of LSAs is questionable. More statistics documenting how students' final standing in senior courses are determined is needed to corroborate this perspective on the utility of the gatekeeping function.

Diagnosis of Classroom Teaching/Instruction

The third purpose of LSAs involves using LSA data for instructional diagnosis and is based on the opportunity for school boards, principals, and teachers to use assessment scores to guide their programs and teaching practices. In this section, the impact of teachers' and principals' skills in interpreting and using LSA data is highlighted as well as issues related to misusing LSA data. Professional development opportunities for teachers to learn about LSAs are also discussed, both broadly and with a specific focus on PE.

Some American researchers purport that LSAs can be used to improve instruction (Landau, Vohs, & Romano, 1999; Popham, 2001). In Canada, Klinger, Volante, and DeLuca (2012) noted that teachers must understand the use of both large-scale and classroom assessment practices, which affect the extent to which they are able to effectively use LSA data to influence their instructional practices. These researchers argued that teachers' shortcomings in assessment literacy provided few opportunities to explore assessment in general, let alone develop expertise in collecting and using data to improve their instructional methods. Stemming from research, there have been appeals for professional development and support to increase the assessment literacy of teachers, administrators, and the community (DePascal, 2003; Volante & Cherubini, 2010). For example, Volante and Cherubini (2010) in their study of secondary teachers' knowledge of Ontario Secondary School Literacy Test (OSSLT), found that teachers' understanding and interpretation of LSA

data came from professional development sessions and staff meetings. In Volante and Cherubini's study, teachers passively engaged with their own assessment scores; moreover, the study reported that provincial data was typically not disaggregated, thereby making it difficult to examine scores from particular student groups or examine relationships between other forms of student data.

In a related study, van Barneveld (2008) also noted that teachers varied in their conceptions of what they considered valuable data and how data should be used. She also suggested that teachers required a clear process of how to examine LSA data, time to acquire skills, and guidance from experts to translate data into useful information. Similarly, Jones (2004) added that "most teachers had not been adequately trained in assessment and needed substantial and ongoing professional development to create valid and reliable tasks and build effective classroom assessment repertoires (p. 586)." Hellsten, Noonan, and Prytula (2013) surveyed 90 principals from two rural and three urban Saskatchewan school jurisdictions on the best way to improve test scores. Professional development in improving assessment strategies was the second most common response from the participants. There were also perceptions that teachers were lacking an understanding of curricula, assessment, and pedagogy, and called for experts in the field to lead professional development on the related topics. Each of these studies illustrates the shortcomings of teachers' skills in interpreting and using LSA data to inform their instructional practices.

There have been a number of other reasons documenting teachers' reluctance to use LSA data to inform their instructional practices. The most common argument was that the feedback from LSAs was too late for schools or teachers to make any changes to their teaching practice (Alberta Assessment Consortium, 2012; Rogers, 2014). Another

researcher who conducted a longitudinal study with 101 participants to examine the culture of data-based decision making in schools reported that teachers' opposition to using LSA results stemmed from idiosyncratic assessment practices and conceptions of teaching that were often resistant to change (Ingram, Louis, & Schroeder, 2004). Similarly, Brown (2004), who surveyed 525 primary school teachers and school leaders in New Zealand, found that teachers associated assessment with school accountability but disagreed with the notion of student accountability. Volante and Cherubini (2010) reflected on teachers' assessment practices and noted that the majority of teachers perceived large-scale external assessments as historically and fundamentally disconnected from their classroom practices. Instead, teachers tended to endorse classroom tests as measures of student performance and, subsequently, drew upon their own test to guide their instruction. The problem with this practice, as noted by Stiggins (2001), was that classroom tests were not reliable instruments and could not provide quality information, nor could they be considered credible enough to hold the system accountable.

Another issue associated with the third purpose of LSAs related to using LSA data to guide teaching can be described as misuses of LSA that involve the practice commonly described as *teaching to the test*. Although some researchers argue that the content standards are not completely aligned with the curriculum in every LSA, the process of creating LSA instruments encompasses practices to ensure that the instrument is aligned with the curriculum (Bhola, Buckendahl, & Impara, 2003; Webb, 1999). Popham (2001) skirted around the issue of teaching to the test, noting that if teachers were effective in teaching the prescribed curriculum, then their students would be successful on LSAs, but more importantly, effective teaching would result in students' mastery of the knowledge or skills

on which the test items were based. Moreover, it was believed that when teachers taught to the test, the curriculum became homogeneous in that what was being taught was the same as the concepts that were being assessed (Kelaghan & Madaus, 1992). In reflecting on this aspect of the issue at hand, it is possible that teaching to the test really is not an issue.

Teaching to the test also raised issues related to teachers' overuse of classroom time to practice for LSAs and drill students' knowledge, all of which reinforced basic skills, such as memorizing, and thereby neglected high-order thinking skills (Sacks, 2000; Volante, 2004). Since LSAs only contained part of the curriculum, teachers tended to focus their instruction to the key areas that may appear on the LSA, which may result in a lack of instruction on other aspects of the curriculum, described by Popham (2001) as *item teaching*. Broadening the argument further, Volante (2004) noted that when teachers focused primarily on the subject being measured, the non-assessed subjects, such as physical education or music, received less attention because instructional time was being devoted to the subjects being measured. Teaching to the test can also lead to weaker and possibly incorrect interpretations about school programs (Mehrens, 1989). As Fay and Smith (2000) noted in their study, teaching to the test can make school results half a year better had they not employed this instructional method. In this context, elevated assessment scores may result in a misunderstanding of students' true ability and, subsequently, a misdirection of funds to other schools with children in need (Volante, 2004).

The last disadvantage of teaching to the test centres on a negative shadow the practice casts on the teaching profession as a whole. As Stiggins (1999) noted, teachers were putting a great deal of pressure on themselves and students as well to do well on LSAs, and teaching to the test only caused the feelings of frustration and disillusionment with the entire

testing process.

Given the use of LSA data to improve instructional practice relies heavily on professional development to adequately prepare teachers to analyze assessment results and use the data in meaningful ways, the PE context was examined for opportunities of professional development in this area. Current teacher training in PE focused on changing short-term behaviours rather than long-term learning needs. The Professional Learning Report (2013) released from the PE Department of Education suggested that professional learning time should be embedded into the school calendar in order to support continuous improvement in teachers' knowledge, skills, dispositions, and practice (PE Department of Education, Early Learning and Culture, 2013). In particular, the document highlighted that professional development was needed to "[identify] procedures to effectively use provincial literacy and numeracy assessment results to assist school board members, school board staff, school leaders, and teachers with data-influenced decision-making (p. 1)." Prior to this initiative, in an interview with the Home and School Federation representative, William Whelan, media had reported problems related to professional development days on PE that needed reform since teachers found it was a waste of time and the effectiveness was questionable at best (Hopper, April 20, 2012). Therefore, it is possible that teachers in PE are not applying their provincial LSA data to guide their instructional development, which then raises concern about the utility of the third function of LSAs.

Although this third function of LSA is theoretically ideal, in reality, the research presented in this section provides evidence of the gap between theory and practice. Subsequently, the utility of the third function of LSA is also questionable.

Monitoring Student Achievement

The last intended function of LSA presented by McEwen (1995) mainly acts as instrument for monitoring student achievement over time. By implementing LSAs in this way, stakeholders can be provided with information on how students are doing in the programs of study. In Canada, as Taylor and Tubianosa (2001) noted, unbiased and objective information is needed to map the trend of student achievement overtime. In Canada, the CMEC clearly emphasized that assessment helps track student progress and achievement, which reflects the function of the three types of LSAs. Take the national PCAP and international PISA for example, both the test development and implementation institutions (i.e., CMEC and OECD) would release the cyclic report regarding student achievement compared to the last round of testing which shows student progress or setback in certain subjects in a three-year cycle.

With regard to provincial assessments, according to DeLuca, Klinger and Miller (2008), twelve out of thirteen provinces across the nation have some form of LSAs that function to monitor student achievement. This approach of comparing performance in a cross-sectional way enables the provincial authority to conduct longitudinal cohort analysis to monitor the same students through Grade 3 to Grade 12 at different key stages. Moreover, they explored that the function of monitoring student achievement is common at the elementary level in order to provide valid information about student learning outcomes.

Due to the lack of relevant research on PE related to this functionality of LSAs, chapter three presents a map of student achievement to monitor the trend by using the provincial assessment data, which provides a glimpse of how students are performing throughout their years in school.

Teachers' Views on LSA

Recent studies have shown that teachers had negative attitudes toward LSAs in North America (Child & Lawson, 2003; Etsey, 1997; Klinger & Rogers, 2011; Kohn, 2001; Rogers, 2014). In contrast to the high-stakes LSAs the United States implemented after the No Child Left Behind Act (2002), LSAs in Canada are regarded as low-stakes (Klinger, Deluca, & Miller, 2008). With no negative consequences to schools or teachers, the accountability framework falls on teachers' professional responsibility with the expectation that school administrators and teachers will use the LSA results to guide and support school improvement. Thus, it is necessary to examine how teachers perceive the implementation, value, and address issues associated with LSAs to garner an understanding of the extent to which teachers use LSA results.

In Klinger and Roger's (2011) study about teachers' perception of LSAs in Ontario and Alberta, they found that teachers from both provinces believed LSA results to be helpful information used to improve learning and achievement of their students. However, teachers' concerns about LSAs included parents' lack of competency when interpreting provincial assessment results, publishing school rankings based on LSA scores, using LSA scores to evaluate teachers, and the manner in which LSA results were reported (Klinger & Rogers, 2011). These concerns contributed to the free information policies in both provinces that enabled public media and agencies (e.g., Fraser Institute, C.D. Howe Institute) to access provincial assessment data. In opposition to LSAs, the Elementary Teachers' Federation of Ontario claimed that teachers knew the students the best, and standardized assessments did not promote accountability because the assessments only provided a snapshot of students' performance (Elementary Teachers' Federation of Ontario, 2015).

Hence, the Elementary Teachers' Federation of Ontario espoused that learning achievement related to the daily meaningful assessments was performed by teachers.

Other teacher federations or associations conveyed concerns related to LSAs. For example, the Manitoba Teachers' Society (2010) claimed that the standardized assessments had no academic value due to its unwieldy information and inability to change on-going curriculum and instruction (2010). Coincidently, the Alberta Teachers' Association (2009) commented that accountability was not about the standardized assessment scores. Rather, they were more in favour of timely classroom assessments that enabled teachers to tailor their instruction to suit student needs and learning styles. The association believed that it is teachers' responsibility to develop measurement strategies to align the curriculum and address students' needs in learning. (Alberta Teachers' Association, 2009). These perceptions of LSAs from teachers demonstrate concerns related to the face validity of LSAs, which is likely to influence the extent to which teachers use LSA findings. If teachers do not believe the LSAs have any value, it is unlikely that they will use the findings to inform their instructional practices, thereby drawing into question the utility of the third function of LSAs to inform teaching and learning.

In 2013, PE Teachers' Federation (PEITF) publicly opposed the budget spent on LSAs because the federation was not satisfied with the provincial assessments (Wright, 2013). The PEITF Union leader, Gilles Arsenault, openly called for the abolishment of LSAs in exchange for hiring more teachers. Arsenault argued that the $1.6 million annually spent in LSAs was not having any impact on the classroom (Zwaagstra, 2014). Drawing on teachers' opposition to LSAs, three of the four candidates for provincial leadership in PE also openly advocated for the abolishment of LSAs (PEI Home and School Federation, 2015).

Similar to teachers' perceptions in other Canadian jurisdictions, PE teachers' position on LSAs influenced the utility of LSAs, and their open opposition to LSAs is likely to influence the public's perceptions of LSAs.

Principals' Views on LSA

The use of LSA data presents an opportunity for principals to demonstrate their leadership, promote teaching quality, and enhance the whole performance of the school. Based on a study by Volante and Cherubini (2010), all of the six secondary school principals interviewed indicated that LSA data constituted one component of student performance, which was considered a starting place to align their annual improvement plan. At the elementary level, 11 principals indicated they were committed to analyzing the data to understand how to link LSA data to the curriculum. Although these researchers noted there were many political factors influencing participants' responses, they believed that principals considered it was important to try to use the data to enhance teaching and learning at the school level.

Another study investigating principals' perceptions on LSAs in Saskatchewan revealed that a large majority (78%) of principals (n=65) indicated that provincial assessments positively influenced teaching and learning in their school. These principals influenced this process by leading the decision making and engaging in priority setting, planning, and classroom instruction (Prytula, Noonan, & Hellsten, 2013). This research echoed another study in Saskatchewan that reported that most principals suggested that the provincial LSA was a key component for school-level planning (Newton & Viczko, 2010). Principals' perceptions on using LSAs to guide teaching and learning is well aligned with the

third function of LSAs that calls for using LSA results to guide instruction and learning. Unfortunately, research examining principals' views on LSAs in PE has not been reported.

Parents' Views on LSA

A small number of studies have examined LSAs from parents' perceptions. In 2004, Osburn, Ritter, Stegman and Suitt (2004) reported that over half (104 of 190) of parents from fifth grade classes in a high-performing school district of Northwest Arkansas considered LSAs to be very important, and over eighty percent (*n*=167) of parents were interested in their child's test scores. However, they were dissatisfied with the absence of communication interpreting the scores. In another parent study by Freeman and Williams (2011), researchers examined minority parents' perceptions of LSAs and reported that parents from culturally diverse groups were more positive towards using LSAs since they believed the standardization process provided a more equitable understanding of students' abilities.

In the Canadian context, Davie, Hart, and Livingstone (2001) found that 46% (*n*=271) of Ontario parents believed that provincial assessments could measure how students are doing and be useful instruments to improve student learning at the elementary level. At the secondary level, 76% of parents supported provincial assessments and were in favour of using assessment results as high school graduation indicators. From the latest Ontario Institute for Studies of Education (OISE) survey released in 2012, less than half of the parents (42%, 42 out of 101) could identify the main role of the Education Quality and Accountability Office (EQAO) (an independent agency that creates and administers provincial LSAs to measure Ontario students' achievement in reading, writing and mathematics at key stages of their education), which was to conduct province-wide

assessments and provide information on student achievement (Hart, 2012). However, 63% of these parents also reported that they had read or heard about province-wide assessment results over the past year since the survey was administrated.

In contrast, another study reported that parents were knowledgeable and supportive of LSAs (Mu & Childs, 2005). In this study, 70% or 104 of parents in Ontario believed that the LSAs provided accurate assessments of individual students, schools, and school boards. Additionally, 72.8% of parents reported understanding standards that guided LSA development and scoring. However, the accuracy of this knowledge could have been validated by asking participants to state one or two of the standards. The majority of parents also agreed that LSAs helped increase accountability of school systems. Mu and Childs (2005) reported that parents wanted more information about other public sources related to LSAs beyond what the school provided. Moreover, parents claimed that the results of LSAs should be more accessible, and sample papers or other resources should be made available to help their child at home. Mu and Childs' study (2005) was fairly positive towards LSAs. However, the other studies described above that followed Mu and Childs' demonstrated a more mediocre awareness of the purpose and value of LSAs possibly because LSAs were fairly new to the Ontario education system, which is likely to influence the accountability function of LSAs given that poor assessment scores may not empower the parental community to raise a call for an improvement in the education system.

Students' Views on LSA

As one of the key stakeholders in the education system, students' perceptions of LSAs also affect the accountability function of LSAs, which has been of interest to some researchers (Klinger & Luce-Kapler, 2007; Miller, 2013). In Klinger and Luce-Kapler's

study (2007), the small but divergent sample of high school students (*n*=42) reported experiencing a great deal of in-class preparation for the Ontario's Secondary School Literacy Test (OSSLT). This preparation for OSSLT led about half of the students to believe that "the test is just like a formula. If [students] put right things in the right place, [they] will pass; [they] do not have to be creative or imaginative at all" (p. 43). These students also reported elevated stressed levels placed on them from their teachers to succeed in the OSSLT.

LSA stress and anxiety has been an issue raised by students, parents, and teachers (Columbia University, 2013; Jones, Jones, Hardin, Chapman, Yarbrough & Davis, 1999). In the study by Jones et al. (1999), 61% of 236 teachers reported that their students were more anxious than normal when preparing for their LSA, and 24% of teachers felt that students were less confident in learning, especially for the low-performing students since their test anxiety was contrasted with successful students. The stress and anxiety concern was conveyed to the public through media reports (Thompson, 2014; Strauss, 2014). This, and other similar information from media reports that influence the level of skepticism towards LSAs, is likely to negatively affect perceptions about LSAs, particularly among parents; however, more research is needed in Canadian contexts.

Among the limited studies on students' LSA perceptions, prior research has shown that high-achieving students often favoured LSAs due to factors related to internal motivation and a strong belief of success. Conversely, low-achieving students were not supportive of LSAs because they knew less about them, and their approach to completing the LSAs was more rudimentary than high-achieving students (Klinger & Luce-Kapler, 2007). It is possible that these students in particular were afflicted by stress and anxiety, given that their knowledge

base was not as robust as their high-achieving counterparts. As previously noted, more research is needed to examine this context.

Public Views on LSA

Current research has shown that public support for LSAs is both consistent and longstanding (Phelps, 2005). A public survey known as the *Public Attitude Toward Education in Ontario* has been administered by the OISE every three years since 1978. Data from the survey provides a snapshot of public opinions about education in Ontario. The latest OISE survey in 2012, which surveyed 1016 adults, revealed that provincial LSAs continued to receive strong support (70% of the participants strongly supported provincial LSAs at the secondary level). However, 67% of participants wanted teachers' assessments rather than LSA scores to determine students' grades. When surveying perceptions related to accountability and LSAs, 64% agreed that testing conducted by the EQAO helps keep the educational system accountable to parents and taxpayers. Although 29% of the public indicated that they had some knowledge about the role of the EQAO in the Ontario education system, 45% of the public reported that they had read or heard about province-wide assessment results over the past year. Almost two-thirds supported province-wide testing in elementary schools in some form, just over half (53%) favoured the current program of measuring every student's achievement each year, and only 11% would supported LSAs that only sampled a small percentage of students. Two-thirds agreed that province-wide testing could be used to improve the quality of education in Ontario. Although Ontario's LSAs enjoy widespread public support, the belief that it has a major impact on achievement is much less widespread. The percentage of the public who thought LSAs had a great deal of impact on student achievement dropped from 50% in 1998 to 40% in 2012 (Hart, 2012).

Periodic monitoring of public opinion about LSAs in Ontario has provided a wealth of information related to the accountability function of LSAs. Although not all participants overwhelmingly supported and valued LSA's influence on the quality of education, there were sufficient responses to support the accountability function in Ontario. There has been no such study like this in the province of PE; hence this thesis aims to provide a vital piece of information about public perceptions on LSAs.

Conclusion

The examination of research related to LSAs has revealed that the utility of the four functions of LSAs varies. A synthesis of this research revealed that the utility of the accountability function of LSAs called for more understanding by all stakeholders as instruments of keeping the education system accountable despite factors compromising issues related to validity. Research from the public's perceptions in Ontario showed that this sample of the population were fairly informed about LSAs, but the types of questions posed to participants did not really check or measure their knowledge. Instead, participants were asked if they knew what the standards were, which solicits different information than asking participants to identify specific standards.

In terms of the gatekeeping purpose, literature has revealed some discrepancies in the practice of using portions of LSA scores to determine students' final standing in a course in Ontario and PE's Grade 9 mathematics classes. In Ontario's Grade 10 LSA of literacy, the repeated attempts to pass the LSA devalued the gatekeeping function. The same is true for PE's LSA Grade 9 mathematics, in which a guideline prevented students from failing a course due to a poor performance on the LSA. Further, there was an absence of research exploring the impact of LSAs to determine students' final grades in the senior grades (i.e.,

grades 11 and 12). Based on the research presented in this thesis, it is possible that the gatekeeping function of LSAs may not be working in the Canadian context.

Lastly, the opposition to LSAs, primarily from teachers, has stymied effective use of LSAs. Teacher opposition has most likely been influenced by a lack of knowledge and time to examine LSA results and apply findings to their instructional practices. In terms of perceptions from varying stakeholders, Ontario parents, students, and taxpayers shared supportive attitudes and favourable values, but whether this perspective is the same in PE is unknown. As previously noted, there has been little research examining perceptions about LSAs, and there has been no research examining the public's perspective of LSAs in PE. This research is critical to understanding the extent to which the purposes LSAs are being met in PE.

CHAPTER THREE

Contextual Framework

The context for this study was situated in Prince Edward Island (PE), the smallest Canadian province, with 146,283 people (Government of PE, 2014a). PE's immigrant population has shown a steady per capita increase in international migrants in comparison to the other provinces[5], having the fifth highest immigration rate in 2014 (9.6% per thousand) (Government of PE, 2015a). From an economic perspective, PE is aligned with the other Atlantic provinces in terms of income and rates of home ownership. The average PE income in 2012 was approximately $69,010, second behind Newfoundland of the Atlantic Provinces (Statistics Canada, 2015). The percentage of PE homeowners was 73.4%, third of the four Atlantic provinces (Statistics Canada, 2013). As for employment opportunities, the unemployment rate in PE was 11.6% as of July 2015, compared with Canada's overall rate of 6.8% (Department of Finance, Prince Edward Island Statistics Bureau, 2015).

Despite relatively comparable living standards and a higher unemployment rate, PE was the worst-performing province in terms of both education and skills, according to the Conference Board of Canada (2015a). This well-recognized organization rated provinces using 23 indicators, measuring educational attainment and skills in the kindergarten to Grade 12 system (drawing on the Program for International Student Assessment results), post-secondary attainment, and also the adult workforce (measured using OECD's Programme for International Assessment of Adult Competencies PIAAC) (Conference Board of Canada, 2015a). These measures demonstrated a "strong and direct relationship between

[5] An exception to this steady increase was in 2012/13 when the provincial nominee program soliciting immigrants was suspended due to an investigation related to misuse (Wright, 2012)

investments in education, educational attainment, and economic growth" (Conference Board of Canada, 2015b). For the province of PE, the situation is not salubrious. This warning of economic hardship warrants an examination of the education system to identify factors influencing the economic wellbeing of the province and its capacity to compete in a globalized world.

Prince Edward Island's Education Program

The two English Language School Boards in PE amalgamated in 2012, resulting in two primary school boards; the English and French (Government of PE, 2015b). During the latter part of the period covered in this study (November 2015), the English Language School Board was dismantled and its functions taken over by the provincial Department of Education because of duplicate services in overseeing curriculum delivery (CBC News, November, 05, 2015). It is also important to note that the PE Department of Education Early Learning and Childhood Development has once again changed its title to the PE Department of Education, Early Learning and Culture. Reference to the Department of Education's documents found on their website reference their new title.

The curricula in PE were developed in accordance with the Canadian national levels (PE Department of Education, Early Learning and Culture, 2013), and are consistent with and reflective of Canadian culture. For example, PE's newly revised Mathematics curriculum was based on the Western and Northern Canadian Protocol, which has been adopted in many Canadian jurisdictions (PE Department of Education, Early Learning and Culture, 2014). Although late entering into the practice of provincial assessments, PE is similar to other provinces in that it participates in LSAs at the international and national

levels and now has its own provincial assessment program. The following sections puts a spot light on PE's student achievement on each of these LSAs.

Prince Edward Island Performance on the PISA

As noted previously, based on the Programme for International Student Assessment's (PISA) 3-year cyclical reports, students in PE have scored at or near the bottom compared with student achievement in other provinces (CMEC, 2001, 2004, 2007, 2010, 2013)[6]. The purpose of this section is to provide a detailed review of PE student achievement in the three key domains commencing in 2000, when the first PISA was introduced in Canada, to the most recent PISA in 2012. The 2015 PISA results were not available at this time of the thesis. This section begins with an overview of the PISA.

When the PISA was implemented in 2000, PE, like other Canadian provinces, was among the first adopters of this international assessment. The 2012 PISA has since been adopted in 65 countries and regions to assess the achievement of 15-year-old students (Grade 10) in reading, mathematics, science, and problem-solving. This age group was chosen because they are nearing the end of their compulsory schooling. Based on the 3-year cyclical reports, students in PE have scored at or near the bottom of the country compared with student achievement in other provinces as previously stated (CMEC, 2001, 2004, 2007, 2010, 2013). Table 3.1 below, ranks PE achievement in comparison to the other nine provincial jurisdictions participating in the PISA (Note: no data were collected in the three territories or on Indian Reserves) (CMEC, 2001, 2004, 2007, 2010, 2013).

[6] The Council of Ministers of Education, Canada (CMEC) publicize detailed provincial reports based on the PISA data in addition to data from the Pan-Canadian Assessment Program which is delivered by the CMEC. Thus references to PISA data may cite the CMEC.

In the PE context, up until the last two rounds in 2012 and 2015, there was a census sample (1,900) due to the small population. Since 2012, students and schools were randomly selected to participant in the PISA. Most 15-year-old students participate in the PISA, except for about 200 students with special education requirements (Department of Education, Early Learning and Culture, 2015b).

Table 3.1

PE Ranking in the PISA from 2000

Year	Discipline	Canadian Rank (out of 10 provinces)
2000	Mathematics	8th
	Reading	8th
	Science	9th
2003	Mathematics	10th
	Reading	10th
	Science	10th
2006	Mathematics	10th
	Reading	10th
	Science	9th
2009	Mathematics	10th
	Reading	10th
	Science	10th
2012	Mathematics	10th
	Reading	10th
	Science	10th

Note. Adapted from CMEC PISA reports 2001, 2004, 2007, 2010, and 2013.

Focus on Mathematics. The PISA assesses students' competency in mathematics content as well as mathematical skills. The four main content areas are change and relationships, space and shape, quantity, uncertainty and data; each of which is equally weighted. In addition, each category is assessed in one of the following contexts: personal, occupational, societal, or scientific. In the 2012 PISA, in which mathematics was the main

domain, the first assessment of students' problem-solving skills was conducted, and in 2012 computer-based testing was also initiated.

Mathematics has been PE students' poorest domain in the PISA. For example, in 2000, PE was 21 points below the Canadian average (CMEC 2001), which is approximately half a grade level behind in mathematics. When considering achievement in mathematics through to 2012, the performance of PE students does not improve; in fact, it gets worse (see Figure 3.1). In 2003, PE was 32 points behind the Canadian average (CMEC 2004). After the release of the Task Force report in 2005 (Kurial, 2005) and what may have been some impetus to improve the mathematics program, PE showed a slight improvement from the previous year, but in the 2006 assessment, PE was still 26 points behind the Canadian average (CMEC, 2007). In 2009, PE was a staggering 40 points behind (CMEC, 2010), and most recently on the 2012 assessment, PE students were 39 points behind the Canadian average, which is more than one full grade level behind in formal mathematics education (CMEC, 2013).

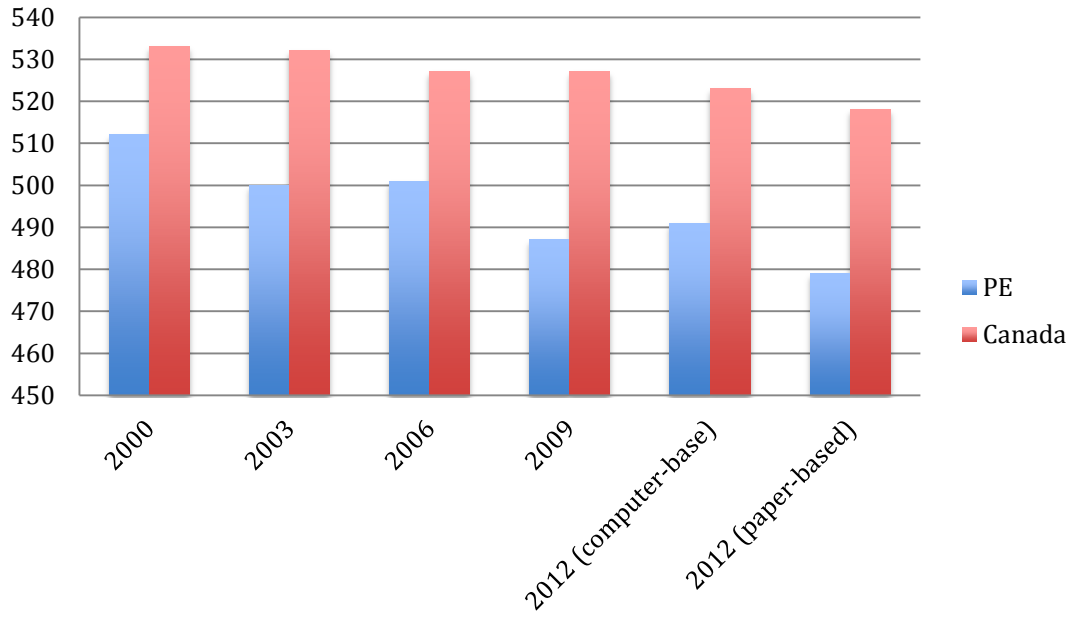


Figure 3.1. PE and Canadian mean scores for student achievement on PISA of mathematics. Adapted from the CMEC PISA reports 2001, 2004, 2007, 2010, 2013.

The PISA measures mathematical proficiency using a 7-point scale, where the easiest tasks are at the lower end of the scale (Level 1 or Below Level 1) and the hardest tasks are at level 6. Level 2 is considered the baseline for mathematical proficiency that is required to fully participate in society. PE had a higher percentage of students who were at the low end and fewer students at the high end of the scale in comparison with other Canadian students. Table 3.2 shows the proportion of students at each level in PE and in Canada. The statistics show data for the 2003 and 2012 PISA, because this grouping of data were only provided for the primary domain being tested, which for the purposes of this discussion is mathematics.

Table 3.2

Distribution of Percentages of PE and Canadian Students by Levels in Composite Mathematics

	PE		CA	
Composite Mathematics	2003	2012	2003	2012
Below Level 1	5.2	4.3	2.4	3.0
Level 1	12.5	14.8	7.7	9.1
Level 2	23.7	28.5	18.3	20.5
Level 3	28.0	31.5	26.2	27.8
Level 4	20.5	17.1	25.1	23.9
Level 5	7.5	3.5	14.8	12.0
Level 6	2.6	0.4	5.5	3.7

Note. Adapted from CMEC PISA reports 2004, 2013.

Focus on Science. The PISA defines science competencies through a combined science scale, which has three main sub-domains of defining scientific issues, explaining phenomena scientifically and using scientific evidences within personal, social, and global contexts (OECD, 2007). During the four cycles of the PISA, science has been the main domain since 2006. According to the Council of Ministers of Education, Canada (CMEC) report that highlighted the PISA trends in Canadian data, students in Newfoundland and Labrador, Prince Edward Island, Nova Scotia, New Brunswick, Manitoba and Saskatchewan performed significantly below the Canadian mean (CMEC, 2007). Similar to mathematics, science achievement was defined by 7 achievement levels, with Level 1 (and below Level 1) as the lowest. The proportion of PE students who performed at the lower level (Level 1 or below) was 16%, which was much higher than the Canadian average of 10% (CMEC, 2007). Conversely, the percentage of PE students with a high proficiency level in science (Level 5 or above) was about 10%, compared with Canadian average of 15% (CMEC, 2007). Once again, PE has the highest number of students performing at Level 1 (and below) and the fewest students performing at Level 5 or higher. Table 3.3 presents PE's

distribution (by percentage) compared to Canadian students' overall science proficiency on the combined science scale which examined scientific literacy, identifying scientific issues, explaining phenomena scientifically, and using scientific evidence.

Table 3.3

Distribution of Percentages of Students at Each Level in Combined Science in PE and that of Canada Overall in 2013 PISA

Combined Science	PE	CA
Below Level 1	3.8	2.2
Level 1	12.1	7.8
Level 2	23.7	19.0
Level 3	29.1	28.8
Level 4	21.5	27.8
Level 5	8.2	12.1
Level 6	1.6	2.4

Note. Adapted from CMEC PISA report 2014.

Focus on Reading. The reading assessment framework in the PISA contains three aspects: Accessing and retrieving, integrating and interpreting, and reflecting and evaluating. Additionally, two text formats of continuous text and non-continuous text were used in the PISA. Although many jurisdictions have improved their performances in reading, PE students continued to consistently rank at the bottom during the four cycles of the PISA between 2000 and 2012. It is important to note that in 2012, the PISA initiated a digital reading version of the assessment. Most Canadian students seemed to perform better on the digital version of the assessment than on the paper version. However, the change of forms did not parallel student achievement in PE. Figure 3.2 compares the Canadian mean and the PE mean in combined reading, which does not show any significant improvement.

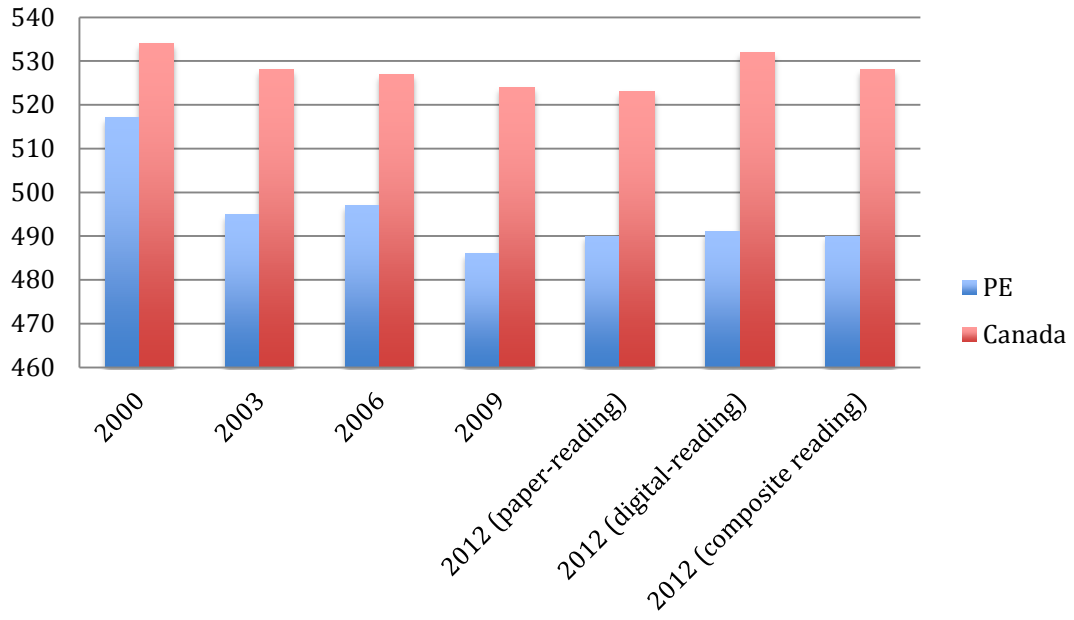


Figure 3.2. PE and Canadian mean scores for student achievement on PISA of reading. Adapted from the CMEC PISA reports 2001, 2004, 2007, 2010, and 2013.

With respect to the three reading sub-domains on the combined reading assessment, PE students once again performed below the Canadian mean in each of these sub-domains in the 2009 PISA. Despite the modification of the reading assessment scale from 5 levels in 2000 to 7 levels (Level 1 was separated into Level 1a and Level 1b) in 2009, PE continued to have the largest proportion of students performing below level 2, which is considered the baseline of proficiency in reading. Table 3.4 shows PE students' achievement in comparison to the Canadian average on the reading assessment in 2009, which revealed that more than 20% of PE students were performing below Level 2, while the proportion of Canadian students performing below Level 2 was only 10%. At the other end of the scale, only 7% of PE students performed at Level 5 or above compared with Canadian mean of 13% (CMEC, 2010). There appears to be a trend of student performance on the PISA of reading where almost double the percent of students scoring Level 1 and below and half the number of PE

students scoring at Level 5 or above. In other words, PE has almost double the percentage of poor performing students and almost half the percentage of high performing students. Table 3.4 shows PE's distribution (by percentage) of students by proficiency level on the combined reading scale compared to the Canadian distribution.

Table 3.4

Percent of Students at Each Level in Combined Reading in PE and that of the Canadian Average in the 2009 PISA

Combined Reading	PE	CA
Below Level 1	1.2	0.4
Level 1	20.0	9.9
Level 2	25.3	20.2
Level 3	27.9	30.0
Level 4	18.7	26.8
Level 5	6.0	11.0
Level 6	0.9	1.8

Note. Adapted from the CMEC PISA report 2010.

Although Canada's mean score in reading performance decreased from 534 in 2000 to 524 in 2009, this was not statistically significant. However, there were five provinces that had significant decreases in reading performance between 2000 to 2009, one of which was PE. PE went from performing above OECD mean, but still ranked the eighth out of the ten provinces in 2000, to performing below OECD average in 2009. Even more concerning was the proportion of low achievers, which increased by 6% from 2000 to 2009 (CMEC, 2010).

Focus on Problem-solving (2012). Since problem-solving are crucial to the well being of individuals as well as society, the OECD began to assess student skills of creativity, interpersonal abilities, project management, and entrepreneurship, which education systems in many countries have recognized as important components in an educational system. Six proficiency levels were defined to score student achievement in problem-solving. PE

student achievement in problem-solving paralleled the poor performance in the previously mentioned domains. On the 2012 PISA, not only did PE students score the lowest but the province contributed the most students at the lowest levels (i.e., Below Level 1 and Level 1) and the fewest number of students at the highest proficiency levels (i.e., Levels 5 and 6). There were large variations in students' performance across the provinces. The span of students reaching baseline proficiency (i.e., Level 2) ranged from 78.8% in PE to 87.5% in British Columbia. For Level 5 and above, the distribution of students' ranged from 7.3% in PE to 19.1% in British Columbia (CMEC, 2014b, p.8). Table 3.5 shows the PE distribution (by percentage) of student proficiency in problem-solving, compared to Canadian distribution in the 2012 PISA.

Table 3.5

Distribution of Percentages of Students at Each Level in Problem-solving in PE and that of Canada Overall in 2012 PISA

Problem-solving	PE	CA
Below Level 1	7.0	5.1
Level 1	14.2	9.6
Level 2	25.7	19.0
Level 3	28.2	25.8
Level 4	17.7	22.9
Level 5	5.6	12.4
Level 6	1.6	5.1

Note. Adapted from CMEC PISA report 2014.

Prince Edward Island Performance in the PCAP

PE students also participate in LSAs at the national level. In 1997, the Canadian Ministers of Education implemented the School Achievement Indicators Program (SAIP), which measured a random selection of schools and students each year between 1997 and 2004. The SAIP measured 13- and 16-year old students in three domains: mathematics,

science, and reading. The SAIP was replaced in 2007 by the Pan-Canadian Assessment Program (PCAP), which similarly measures a random selection of schools every three years but only assessed 13-year old students (CMEC, 2008).

Like the PISA, the PCAP is not an indicator of individual student achievement but rather an instrument to compare student achievement at the provincial/territorial level as well as to provide feedback for curriculum and provincial assessment development (CMEC, 2014). Table 3.6 below highlights PE student achievement compared to the Canadian mean in the three areas of assessment in the PCAP from 2007 to 2013.

Table 3.6

PE PCAP Rankings and Mean Scores Compared with Canadian Scores in Mathematics, Science and Reading From 2007 to 2013

	PE Ranking	PE Mean Score	Canadian Mean Score
2007			
Mathematics	11th	450	500
Science	10th	464	500
Reading*	11th	460	500
2010			
Mathematics*	11th	460	500
Science	4th	493	500
Reading	8th	481	500
2013			
Mathematics*	4th	492	507
Science	6th	491	500
Reading	6th	494	508

Note. * Indicates the main domain in each cycle of PCAP. Jurisdictions that took part in PCAP from 2007 to 2010 are BC, AB, SK, MB, ON, QC, NB, NS, PE, NL and YK. In 2013, there was only 10 provinces/territories paticipanted in PCAP, since the YK withdrew.

Mathematics Performance. Slightly different from the PISA, the framework for the PCAP was based on the National Council of Teachers of Mathematics (NCTM) principles and standards for school mathematics, which most of the jurisdictions currently use for mathematics curriculum design (CMEC, 2011). Similar to the PISA of

mathematics, the PCAP measures students' mathematics competency in four domains: numbers and operations, geometry and measurement, patterns and relationships, and data management.

PE students experienced a large improvement from 2010 to 2013. On the first two assessments, students were situated at the bottom of the national scale of the eleven provinces. In the most recent mathematics assessment (PCAP 2013), PE jumped to the fourth position, with a mean score of 492 compared to the Canadian mean of 507, which was considered a significant improvement in mathematics.

The PCAP of mathematics also uses levels to define students' mathematics achievement using a four-point scale to measure student proficiency. Low cognitive tasks are placed at the lower end of Level 1, while sophisticated mathematics problems are grouped at Level 4. Based on the curriculum expectations in mathematics, students are required to achieve the basic level of 2 or above. At Level 2, students can recall facts, make definitions, perform operations, evaluate expressions, and retrieve information. Typically, students can solve problems that are cognitively low or moderate (CMEC, 2011). Table 3.7 shows the proportion of students at each level in PE and that of the Canadian sample in the 2010 PCAP of mathematics. Compared with the percentage of students at each level, there was a clear trend that PE had more students (70%) situated at the lower end of the achievement (i.e., Level 1 and Level 2) compared with Canadian students overall (54%). In contrast, only 29% of PE students scored at the higher level of mathematics performance (i.e., Level 3 and Level 4), much lower than that of the overall Canadian average of 47%. It should be noted that the percentage of PE students at each level on mathematics assessment

was not presented in the reports of the 2013 PCAP, where the large progress was seen given that mathematics was not the main domain of the 2013 PCAP.

Table 3.7

Percent of Students at Each Level of Mathematics Achievement in PE and that of the Canadian Overall in the 2010 PCAP

	PE	Canada
Level 1	13	9
Level 2	57	45
Level 3	29	43
Level 4	0	4

Note. Adapted from CMEC PCAP report 2011. Percentage at each level was calculated only the subject was the main domain in PCAP assessment.

Science Performance. The CMEC adopted a similar definition of scientific literacy as the OECD had in its PISA to enhance the possibility of comparison between the two assessments. The PCAP science assessment uses a four-point scale to measure students in three competencies (science inquiry, problem solving, and scientific reasoning), and four subdomains (nature of science, life science, physical science, and earth science).

PE students did not do well on the first science assessment in 2007, with its ranking of ten out of eleven provinces/territories. The mean score (464) was significantly lower than the Canadian mean score of 500. In the 2010 PCAP, there was a large improvement in science, which placed PE in the fourth position out of the eleven jurisdictions across the nation. PE students' mean score was at the same level of the Canadian overall. The recent PCAP (2013) reported that the mean score for PE students was 491 in science, which was significantly lower than the Canadian average of 500; a drop from the previous LSA. In terms of the proportion of students at each of the four levels, PE had 57% of students at the lower level (i.e., Level 2 or below), while the percentage of students at the higher level (i.e.,

Level 4) was 2% less than the Canadian overall average. Table 3.8 presents the detailed information about this distribution.

Table 3.8
Percent of Students at Each Level of Science Achievement in PE and that of the Canadian Overall in the 2013 PCAP

	PE	Canada
Level 1	7	8
Level 2	50	44
Level 3	37	39
Level 4	6	8

Note. Adapted from the CMEC PCAP report 2014.

Reading Performance. Reading was the main domain of the first PCAP in 2007. Three subdomains of reading were assessed: comprehension, interpretation, and response to text. Student proficiency in reading was presented in three levels of performance, where Level 2 was designated as being the acceptable level of performance for Grade 8 students. Students who achieved Level 2 are expected to comprehend, interpret, and respond to the text, draw conclusions, and demonstrate their knowledge of organizing texts according to different purposes. Level 3, then, represents the highest achievement where students are able to answer the most challenging questions as well as the questions in Level 1 and Level 2 (CMEC, 2008).

In terms of ranking and the mean scores of reading assessment for the 2007 PCAP, PE students scored at the lowest position of the eleven provinces/territories with the mean score of 460, which was significantly lower than Canadian mean of 500. Table 3.9 shows the percent of students at each level in PE and that of the Canadian overall. In the subdomains of the reading test, the mean score for PE students was significantly lower in the interpretation and text response. In the 2010 PCAP, PE ranked eighth out of the eleven jurisdictions, with a mean score of 481 compared to the Canadian mean score of 500. In the

latest cycle of reading tests in 2013, PE ranked sixth, with an average score of 494; however, this reading achievement was still significantly lower than the Canadian mean score of 508.

Table 3.9

Percent of Students at Each Level of Reading Proficiency in PE and that of the Canadian Overall in the 2013 PCAP

	PE	Canada
Level 1	19	12
Level 2	70	66
Level 3	11	22

Note. Adapted from the CMEC PCAP report 2008.

Generally, gender differences in achievement scores occured in the assessment of mathematics and science. In the 2007 PCAP, reading achievement for both girls and boys in PE was significantly lower in comparison to the Canadian mean. In the 2007 PCAP of reading, the mean score of PE girls was 481, while the mean score of boys was 458; a 23 point gap. Girls outperformed boys in reading both within the province (by 30 points) and in Canada overall (by 27 points) (CMEC, 2014).

Prince Edward Island Achievement Trends on the PISA and the PCAP

Based on the national and international LSA results, a general trend can be mapped in the main testing domains. For example, of the group of Grade 8 students who took part in the 2007 PCAP, this population of students also wrote the PISA test two years later; albeit the students may not have been the same students but the population would have been the same. As can be seen from the test results from these two LSAs, the reading performance of PE students did not improve from their last place compared to other jurisdictions nation-wide. It is important to note that before 2009, there was no reading assessment at the provincial level for PE students in Grade 8 and then again in Grade 10. Similar comparison can be made between the mathematics assessment of the 2010 PCAP and the 2012 PISA.

In the 2010 PCAP, PE students were ranked at the bottom, a situation that did not change two years later when this population of students wrote the PISA in 2012. At that time, the province had been administering the Grade 9 mathematics assessment for five years from 2007 to 2012 for the purpose of instructional development and student diagnosis. Given the purpose of this third LSA, one would expect students' achievement in mathematics to improve which should have been seen on the national and international assessments. Although the same cohort of students who took both assessments may not be identical, they are sufficiently close due to the census sampling, to allow for inferences to be made about the population and in this case, it is reasonable to claim that PE students did not make any progress in the key testing domains.

Indicators that Affect Student Performance in the PISA and PCAP

The PISA and PCAP also administer questionnaires to gather information about students' family background, home environment, and educational support as part of each LSA. A separate questionnaire was distributed to school principals to obtain information about principals' and teachers' perceptions and practice of using LSA data to improve the school's overall performance.

Both the PISA and PCAP reports identify three main factors as influencing student performance in LSA: family background, socioeconomic status, and educational support. Family background mainly refers to family structure (single parent or two parents) and whether the student is from an immigrant or non-immigrant family. Socio-economic status focuses on family possessions, home educational resources, and number of books at home. Educational support also has a number of indicators, such as parents educational background, parental involvement, and expectations. These factors were investigated in relation to

student achievement in the three domains. The findings presented below are drawn from the 2000, 2003 and 2006 PISA reports because the PCAP's reports are largely drawn from the PISA's data.

Family background. Family structure divides students into two categories: those in single-parent families and those in two-parent families. In half of the 14 countries examined, including Canada, students from two-parent families had significantly higher levels of achievement than did students from single-parent families in reading. Provinces that contributed to this statistic in reading were New Brunswick, Manitoba, Saskatchewan, Alberta and British Columbia. As for PE, while there was a small gap between the number of single family (503 points) and two-parent family (510 points), student achievement from two-parent families was not significantly higher than single parent families (CMEC, 2001, p69). It is not clear from the available data why this was the case. No evidence of potential influences from family background would affect student achievement on mathematics and science on the PISA.

Immigrant status. When focusing on immigrant characteristics, the PISA identifies two groups: first-generation immigrants[7] and second-generation immigrants[8]. Based on the overall performance in science in the 2006 PISA, non-immigrant students outperformed second-generation immigrant students even though they had similar school experiences in Canada. At the same time, difference in scores in science performance between students of non-immigrants and first-generation immigrants was significant in Quebec, British Columbia and Manitoba (CMEC, 2007). However, since data from PE were not presented in the PISA report, we are unable to draw conclusions about PE students' science performance. The

[7] First-generation immigrant students are those who were born in Canada.
[8] Second-generation immigrants students are those whose parents born in Canada.

possible reason for that may due to the small sample of immigrant students from PE. Take the example of the 2010 PCAP assessment, the questionnaire reported that there was only 16 students in PE who self-identified that they were not born in Canada, which constituted 3.4% of the total population of 1,655 students who took the test (M. Maclean, personal communication, November 9, 2015). Thus it is reasonable to assume that due to the small number of immigrant students, PE students' performance on national and international LSAs had not been affected by the number of immigrant students.

Parents educational attainment. The relationship between parental education and student performance was examined with respect to the 2003 PISA (mathematics as the major domain). In Canada, very few parents were reported to have less than high school education. Therefore, the lowest four categories were grouped together as "high school or less." The higher levels of parental education were defined as college diploma or university degree. As a whole, the mean score of Canadian students whose parents had high school or less was 515, while the mean score of students whose parents had college or university education were respectively, 531 and 553. Interestingly, the PE context did not follow this trend since nearly 40% of parents have a university diploma (a higher percentage than in some provinces, such as Quebec at 35%) but PE students nevertheless had the worst performance (500 in combined mathematics) compared with students from other provinces (CMEC, 2004).

After analyzing parental education from the 2003 PISA survey, the CMEC concluded noting that high parental education was associated with higher mathematics performance (CMEC, 2007, p53). This trend was echoed in the International Standard Classification of Occupation (ISCO), which also acknowledged that parental occupations and their related job

skills were associated with student achievement in mathematics (CMEC, 2004). Although parental education was not a decisive or sole factor determining student performance in mathematics, the 2006 PISA report documented that students whose parents had a university degree tended to perform better than those whose parents had high school or less. Since parental education in PE was quite high in comparison to other provinces, with fewer students whose parents with high school or below educational attainment combined with the large number of students whose parents with post-secondary education (CMEC, 2007, p85), the trend identified by the CMEC based on the PISA data would suggest that PE students should have high achievement in mathematics. However, this was not the case as previously demonstrated in the above section.

Other than the student background factors identified in the questionnaire accompanying the PISA and the PCAP, perceptions and practice about using LSA data were also examined on the questionnaire given to school administrators and teachers. In the 2010 PCAP database containing PE principal surveys for each PCAP administration, over 80% of 465 school administrators in PE strongly agreed (23, 4.8%) or agreed (379, 78.5%) that the PISA/PCAP results were easily obtained. However, just over half the of the principals (54.3%) thought that the results were easily interpreted. Sadly, only 16.1% of the 458 school administrators agreed or strongly agreed that the PISA/PCAP results were easy to use in making instructional change. When surveying school administrators about communicating LSA results to parents or guardians, 55.3% of the 466 sample population disagreed or strongly disagreed with this action. These data illustrate the low value that PE principals/school administrators held for national and international LSAs. Thus the

negative attitude may influence the public perceptions about LSAs, especially for the national and international assessments.

Prince Edward Island Provincial Assessment

The province of PE administers several LSAs at key stages of student development. These include the Early Years Evaluation of four-year-olds beginning formal schooling; the Grade 3, 6, and 9 Language Arts and Mathematics assessments; and most recently PE has piloted a Grade 10 Literacy and Grade 11 Mathematics assessment. These are in addition to the province's participation in national and international LSAs (i.e., the PCAP and PISA). The data discussed in the following sections is drawn from several sources: the public data from the website of the PE Department of Education, Early Learning and Culture; the public data released in the Annual Report from the Department of Education of Prince Edward Island; and the data from the Instructional Development and Achievement Office of the PE Department of Education, Early Learning and Culture by personal contact. A major challenge in analyzing these data was the inconsistency from year to year and the absence of information provided in the annual report that is publicly available on the Department of Education's website. In order to mitigate this difficulty, I have compared as many sources as possible and confirmed the data with authorized persons in the relevant departments including: specialists from the Instructional Development and Achievement Office of Department of Education, local high school teachers, and administers.

The Early Years Evaluation was first administered in 2010 to four-year-old children before they began public school in PE (PE Department of Education, Early Learning and Culture, 2012). This assessment contained two parts: province-wide assessment of all children before kindergarten and classroom based assessment completed during the

kindergarten year by the classroom teacher (Boyd, 2010). The province-wide assessment takes 30-minutes for the purpose of identifying possible academic or developmental delays. Age-appropriate tasks would be asked by the assessors (i.e., teachers trained at the primary level) to measure students' abilities in five main domains: awareness of self and environment, social skills and behaviour and approaches to learning, cognitive skills, language and communication, and physical development. Individual students' outcomes of this part of the assessment were shared with the parents and the child's kindergarten teacher but not the public. The second part of the early years evaluation is a ten-minute observational survey performed by the kindergarten teacher. Data from this assessment is also not publicly available nor is any information on the extent to which the data is being used for its purported diagnostic purposes; despite being referred to as an "evaluation," a term reserved for the critique of programs, whereas a critique of students' abilities is most often referred to as "assessment" (Hodnett, 2001).

PE's LSA program for Grades 3, 6, and 9 was first introduced in 2007 and was the last educational jurisdiction to implement provincial LSAs (Klinger, DeLuca, & Miller, 2008). The Grade 3 assessment in Language Arts (Primary Literacy Assessment) and Grade 9 assessment in Mathematics (Intermediate Mathematics Assessment) were first introduced in 2007, followed by the Grade 6 Language Arts assessment (Elementary Literacy Assessment) in 2008 and Grade 3 Mathematics assessment (Primary Mathematics Assessment) in 2009. Grade 9 Language Arts (Intermediate Literacy Assessment) and Grade 6 Mathematics (Elementary Mathematics Assessment) were implemented in the 2012-2013 school year. Annual provincial assessments are now conducted in Language Arts and Mathematics for Grades 3, 6, and 9.

In January 2015, the first Grade 11 assessment of mathematics, known as Secondary Mathematics Assessment, was piloted. The purpose of this LSA was to ensure students have acquired enough mathematical knowledge and skills for the remaining years before they graduate from high school. This mandatory provincial assessment will be worth 25% of students' overall mark in the course and will be recorded on students' transcript and report card. The official launch of the Grade 11 mathematics assessment will occur in January 2016, which is a paper pencil based assessment.

Also beginning in 2016, PE students will complete an additional Grade 10 literacy assessment (Secondary Literacy Assessment), which will be regarded as a graduation *indicator* until 2019 when it will be considered a graduation *component* (The government of Prince Edward Island, 2015). This LSA is similar in design and administrations to other provinces (e.g., Ontario). In fact, PE's Grade 10 literacy assessment was borrowed from Ontario's LSA testing body (the Educational Quality and Assessment Office, EQAO) and will be administered via a computer, which is a new forum for LSA delivery in PE. This LSA can be described as "moderate stake," because if students fail the test in Grade 10, the government promises intervention strategies and an additional chance to write the test in Grade 11 (and again in Grade 12 if students are persistently unsuccessful) as described in the previous chapter. Students must successfully pass the test in order to graduate from high school. Table 3.10 below, highlights the year, grade, and subject each LSA was introduced or has been proposed in PE compared to Ontario.

Table 3.10

Year, Grade, and Subject each LSA was Introduced or has been proposed in PE Compared to Ontario

Grade	Subject	Year of Implementation	
		PE	ON
K	Early Years Evaluation (EYE)*	2010	n/a
3	Language Arts*	2007	1996-1997
3	Mathematics*	2009-2010	1996-1997
6	Language Art	2008-2009	1998-1999
6	Mathematics	2012-2013	1998-1999
9	Mathematics	2007	2001
9	Language Art	2012-2013	n/a
11	Mathematics	2015 (January)	n/a
10	Literacy Test*	2016	2002

Note. *EYE program is a universal screening program intended to evaluate children's ability before they enter kindergarten.
* Language Arts refers to the assessments in reading comprehension and writing.
* Grade 3 Mathematics were only administrated in Western School Board in 2009.
* Grade 10 Literacy test for high school graduation requirement will be for 2019 cohort.

PE Student Performance in Provincial Assessments

Grade 3 Reading and Writing. Grade 3 Language Arts assessment, known as the Primary Literacy Assessment, was the first provincial LSA developed and implemented in 2007. The writing assessment contains two subdomains of personal expressive and transactional writing, which are assessed separately. Based on reports obtained from the Instructional Development and Achievement Office of the Department of Education, there has been a consistent improvement in reading, from 62% of students meeting or approaching the standard in 2007 to 88% in 2013 (English program). The percentage of students who experienced difficulties dropped from 27% in the first assessment (2007) to 8% in 2012. Unfortunately, data summarizing student performance in 2014 and 2015 was not included on the provincial assessment reports from the Instructional Development and Achievement Office. Hence it is not possible to explore the extent of this trend currently.

For the first cycle of the writing assessment, from 2007 to 2009, the data obtained

from the Instructional Development and Achievement Office presented the percentage of students who both met the standard and who failed to meet the standard with comprehensive results from the two subdomains of personal expressive and transactional writing. From 2010 to 2013, the reports only showed the percentage of students who achieved the standards in the separate subdomains. These data showed that approximately half of the students met the expected level in the first three years in both of the writing subdomains, with 56% in 2007, 45% in 2008, and 53% in 2009. In the personal expressive component, there was a range from 74% to 79% of the students meeting the standard from 2010 to 2012, which was followed by a sharp drop to 62% in 2013. However, for transactional writing part, which was considered to be a challenge for PE students, a slightly lower percentage of students achieved the standard than in the personal expression, with 68% in 2010, 71% in 2011, 67% in 2012, with a gain to 71% in 2013. Table 3.11 and Table 3.12 present the results from Grade 3 Language Arts assessment in reading and writing respectively.

Table 3.11

Percent of PE Students at Each Level of Achievement on the Grade 3 Reading Component of the Language Arts Assessment (Primary Literacy Assessment) from 2007 to 2013 (English Program)

	Met Expectations	Approached Expectations	Experienced Difficulty
2007	62	11	27
2008	72	8	20
2009	85	5	10
2010	82	n/a	n/a
2011	85	5	10
2012	88	5	8
2013	88	n/a	n/a

Note. Adapted from provincial assessment reports from Instructional Development & Achievement Office of Department of Education of PE.

Table 3.12

Percent of PE Students at Each Level of Achievement on the Grade 3 Writing Component of the Language Arts Assessment (Primary Literacy Assessment) from 2007 to 2013 (English Program)

	2007	2008	2009	2010	2011	2012	2013
Personal expressive							
Met expectations	67	56	60	74	79	79	62
Not met expectations	33	44	40	n/a	n/a	n/a	n/a
Transactional							
Met expectations	92	73	71	68	71	67	71
Not met expectations	8	27	29	n/a	n/a	n/a	n/a
Both exercises							
Met expectations	56	45	53	n/a	63	62	n/a
Not met expectations	44	55	47	n/a	n/a	n/a	n/a

Note. Adapted from provincial assessment reports from Instructional Development & Achievement Office of Department of Education of PE.

Grade 6 Reading and Writing. The Grade 6 Language Arts assessment (Elementary Literacy Assessment) was first implemented in 2008-2009 school year with a reading test for students in the English program (an examination of the French Language Arts program is outside the scope of this study). The writing assessment was added in the 2009-2010 school year. Similar to the Grade 3 Language Arts assessment, the writing assessment contains two subdomains of personal expressive and transactional writing, which are assessed separately. The percentage of students who were at or approached the standard in reading has varied through the 7 years (from 2008 to 2015), with 71% of students meeting the standard in 2008, followed by a sharp drop to only 52%, which was attributed to students' lack of skills in non-fiction reading (Department of Education and Early Childhood Development, 2010), to a height of 78% in 2012.

As for the writing assessment, there had been a stable percentage of students at the standard in personal expressive subdomain, with 72% in 2011 and 79% in 2012. Then, the number dropped to 75% in 2013 according to the newly released provincial assessment

results from the website of Department of Education in 2015. With respect to the transactional writing component, the results provided by the Instructional Development and Achievement Office was only available for the years spanning 2009 to 2013. There was no obvious growth in the percentage of students who met the standard in the first two rounds of the assessments, with 74% and 73% in 2009 and 2010 respectively. Then there was a slight drop to 66% in 2011, followed by a dramatic increase to 77% in 2013. Table 3.13 and Table 3.14 show detailed information about students' achievement on Grade 6 Language Arts assessment.

Table 3.13

Percent of PE Students at Each Level of Achievement in Reading on Grade 6 Language Arts Assessment (Elementary Literacy Assessment) from 2008 to 2013 (English Program)

	Met Expectations	Approached Expectations	Experienced Difficulty
2008	71	9	20
2009	52	11	37
2010	66	n/a	n/a
2011	66	n/a	n/a
2012	78	n/a	n/a
2013	76	n/a	n/a

Note. Adapted from provincial assessment reports from Instructional Development & Achievement Office of the Department of Education of PE.

Table 3.14

Percent of PE Students at Each Level of Achievement in Writing on Grade 6 Language Arts Assessment (Elementary Literacy Assessment) from 2009 to 2013 (English Program)

	Personal Expressive		Transactional	
	Met Expectations	Not yet Met Expectations	Met Expectations	Not yet Met Expectations
2009	74	26	74	26
2010	76	n/a	73	n/a
2011	72	n/a	66	n/a
2012	79	n/a	77	n/a
2013	75	n/a	-	-

Note. Adapted from provincial assessment reports from Instructional Development & Achievement Office of the Department of Education of PE.

Grade 9 Reading and Writing. The Grade 9 Language Arts assessment (Intermediate Literacy Assessment) is considered new, since it has only been administrated to PE students in the English Program since the 2012-2013 school year. This provincial assessment contributes to 10% of students' final standing in the course (Department of Education and Early Childhood Development of Prince Edward Island, 2015). The result of the first test from the annual report showed that there were 79% of the students at or approaching the standard in reading (PE Department of Education, Early Learning and Culture, 2015f). In fact, the Instructional Department and Achievement Office of the Department of Education released the data showing that 72% students had actually met the expectation while the remaining 7% of students had not met the standard but *approached the standard*. This percentage increased a little from 79% to 81% in the second assessment in 2014. A huge gap was found in the writing assessment compared to the year before, with 69% students meeting or approaching the standard in 2013, while in the annual report for 2013-2014, the percentage of students who were at or approached the standard had decreased

to 59%. Table 3.15 below presents the results of Grade 9 Language Arts assessment in 2013.

Table 3.15

Percent of PE Students at Each Level of Achievement on Grade 9 Language Arts Assessment (Intermediate Literacy Assessment) in 2013 (English Program)

Writing		Reading		
Met Expectations	Average Score	Met Expectations	Approached Expectations	Experienced Difficulty
69	83	72	7	21

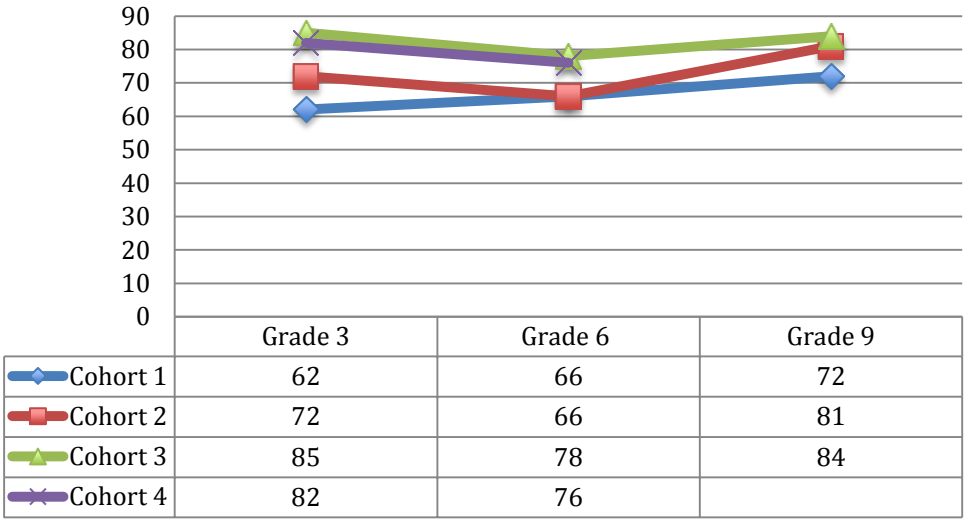
Note. Adapted from provincial assessment reports from Instructional Development & Achievement Office of the Department of Education of PE.

Based on the data from these provincial assessments, the achievements of four cohorts of students can be traced to map students' literacy achievement. Taking the reading test for example, starting in 2007, 62% of the first cohort of Grade 3 students took the Language Arts assessment and met the expectation level. Three years later, in 2010, when these same students were now in Grade 6, the result showed that 66% of students met the standard. When they were in Grade 9 in 2013, the percent of students who were at the standard increased to 72%.

The second cohort of Grade 3 students took the assessment in 2008, with 72% of students who met the standard, however, this number dropped to 66% when they wrote the Grade 6 assessment in 2011, then there was a huge increase in the number to 81% of students meeting the standard or approaching it when this cohort of students moved to Grade 9. It is important to note that this percentage of 81% does not represent the actual number of students meeting the standard, but a sum of students at and approaching to the standard; hence this percentage is inflated and raises issues with the ethical implications of reporting practices in PE.

The third cohort of students began their Grade 3 reading assessment in 2009, with 85% of students achieving the standard. However, the result of the Grade 6 assessment showed that only 78% of students met the standard three years later; a drop of 7%. When students moved to Grade 9, 84% of students met or approached the standard. While this percentage seems to have increased, it should be cautioned that this number does not reflect the actual number of students who met the standard, but includes both that number and those only approaching the standard. Thus, it is difficult to tell whether there was any improvement or drop in reading in this cohort of students from Grade 6 to Grade 9, but a decrease in performance in Grade 6 assessment from the previous assessment results in Grade 3 was noted.

The fourth cohort students wrote Grade 3 reading assessment in 2010, with 82% of them meeting the expected level. Unfortunately, this number dropped to 76% in 2013 when they took the Grade 6 assessment. For this cohort, the Grade 9 assessment, will take place in 2016. As seen in the above analysis, PE students' reading ability does not seem to improve but rather to gets worse between Grade 4 and Grade 6, followed by a slight progress through the subsequent three years to Grade 9. Figure 3.3 presents the percentage of students who met or approached the standard in reading on the provincial assessment from Grade 3 to Grade 9 in four cohorts.



	Grade 3	Grade 6	Grade 9
Cohort 1	62	66	72
Cohort 2	72	66	81
Cohort 3	85	78	84
Cohort 4	82	76	

Figure 1.3. Percent of Students who Met or Approached the Standards in Reading on the Provincial Assessments of Four Cohorts (English Program). *Note.* The results of the second and the third cohort of students in Grade 9 reading assessment were the percentage of students met or approached the standards, the rest of the data presents the actual percentage of students met the standards on provincial assessments. Adapted from the provincial assessment reports from Instructional Development & Achievement Office of the Department of Education of PE.

The same first cohort of Grade 3 students who wrote the writing assessment in 2007, at which time only 56% of students met the standard in both exercises of personal expressive and transactional subdomains (67% of students met the standard of personal expressive and 80% of the students met the standards of transactional). Considerable improvement was made when they took the Grade 6 writing test in 2010, with 76% meeting the standard in personal expressive and 73% in transactional. However, three years later (2013), the Grade 9 writing assessment showed that there was a drop to 69% of students achieving the standard, which there was no subdomain data to be obtained.

A similar trend can be found in the second cohort of students, who entered Grade 3 in 2008. While only 45% of students met the standard in personal expressive, and 55% in transactional writing in Grade 3, huge progress was shown when they wrote the Grade 6

writing assessment (2011), with 72% who met the standard in personal expressive and 66% in transactional writing. Echoing the first cohort's results, there was a sharp decrease to only 59% meeting the standard when they wrote the Grade 9 assessment.

A slight but steady improvement was seen in the third cohort of students, who had their Grade 3 writing assessment in 2009. Only 53% of students met the standard in personal expressive and 47% in transactional writing in the first test, but the percentage of students who achieved the standard rose dramatically to 79% and 77% respectively in Grade 6; followed by a more consistent result in Grade 9 (in 2015), with 78% of students meeting or approaching the standard.

Less progress was seen in the fourth cohort, with 74% of students meeting the standard in personal expressive and 68% in transactional in Grade 3, compared to 75% of students meeting the standard in Grade 6. However, they began at a much higher level than the second and third cohorts. Table 3.16 presents the percentage of students who met the standard in writing on the provincial assessments in Grades 3, 6, and 9 which demonstrates the trend of writing achievement from 2007 to 2015.

Table 3.16

Percentage of PE Students Met the Standard in Writing Results of Provincial Assessments of the Four Cohorts (English Program)

	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Grade 3				
Personal expressive	67	56	60	74
Transactional	80	67	73	68
Grade 6				
Personal expressive	76	72	79	75
Transactional	73	66	77	76
Grade 9				
Total score	69	59	78*	n/a

Note. Grade 9 writing assessments only provide combined results. Adapted from provincial assessment reports from Instructional Development & Achievement Office of the Department of Education of PE (2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014). * The Grade 9 writing score of Cohort 3 was adapted from "Provincial Assessment Results 2015-Backgrounder," released from the website of Department of Education, Early Learning and Culture, 2015.

In conclusion the review of PE's LSA of Language Arts in reading, there has been a slight drop in the percentage of students who met or approached the standard in Grade 6 in cohorts 2, 3 and 4. From the trend observed in Figure 3.3 above, there has been no progress in reading through Grade 4 and Grade 5. With respect to writing, students had better skills in the personal expressive form than transactional form of writing. Furthermore, a drop in the percentage of students who met the standard between Grade 6 and Grade 9 assessments can be identified in the first two cohorts, but this seems less obvious in the third cohort and at this time, there was no data for the fourth cohort. No progress appears to have been made in writing skills between Grade 7 and Grade 8.

An interesting trend can be found when comparing Ontario and PE provincial data. According to the EQAO, the percentage of students who were at or above the provincial standard in reading proficiency in Grade 3 had ranged from 62% to 70% through 2010 to 2014. Unlike PE, much progress was made in grade 4 and 5 in Ontario schools such that the percentage of students who met or exceeded the reading standard increased from 72% to

79% in Grade 6. The Ontario Secondary School Literacy Test (OSSLT) showed that the percentage of students who successfully met the standard was stable during 2011 to 2015, around 82% to 83%. An increase in the percent of successful students demonstrated a positive trend of student progress in reading proficiency, which may explain why Ontario students performed well enough to be in the top two jurisdictions in the PCAP of reading through 2007 to 2013. In terms of PE, conversely to Ontario, the percent of students at or approaching the standard in reading was actually decreasing when comparing data from Grade 3 and 6. No obvious progress can be traced both in reading and writing achievement in the cohorts of students mentioned above. Since there is no clear and detailed explanation of the standards set in provincial assessments in PE, even though the percent of students who met expectations is not drastically low, it is important to note that from the available data, students are generally not doing particularly well in school according to the PISA. In PE, students appear to be struggling which is inferred from mapping student achievement through key grades and drawing on the national and international LSA reports of student achievement.

It is also important to note that since 2014, the PE Department of Education has started a new system of reporting students' performance from provincial assessments. Prior to 2014, public data could be found in the annual reports posted on the website of PE Department of Education. The annual reports present provincial data in terms of the percentage of students who met or approached the standard in reading, writing, and mathematics in Grade 3, 6, and 9 on provincial assessments. The new reporting system is mainly based on online and can be adapted to select individual school or school board data. There is no tool to compare school-based data or track cohorts of students from one

assessment to the next. Government publications related to provincial LSAs usually contain year to year data which can track how students are doing in key subjects and provide parents, educators, and policy makers with a general idea of students' strengths and weaknesses.

Compared to Ontario, PE provincial assessment data does not appear to be utilized fully or appropriately. Public data obtained from the website was not accurate in communicating the actual percentage of students who met the expectations on each of the assessments in Grade 3, 6, and 9. Detailed and accurate information was not easily obtained and no specific research reports have been provided to the public in terms of student achievement throughout the years of LSAs. This difficulty in accessing the data combined with inaccurate data makes it difficult for the public to accurately understand students' LSA in this province. Subsequently, poor reporting practice may hinder participants' knowledge about students' LSA achievement and possibly their perceptions about LSA which in turn may affect the accountability function of LSAs.

Trends in Grades 3, 6 and 9 Mathematics

Grade 3 Mathematics. Prince Edward Island's Grade 3 mathematics assessment (Primary Mathematics Assessment) was officially introduced in 2009 and was only administered in the Western School Board. In 2010 it was applied to all school boards across the province. Unlike the literacy data, the results of the mathematics assessments are available for both the English and French language school boards.

In the Grade 3 mathematics assessments from 2010 to 2013, the average score was quite stable, ranging from 75% to 80%. According to the annual report from the PE Department of Education (2014), 83% of the students were at or approaching the standard. Data from the Instructional Development and Achievement Office of the Department of

Education showed that the actual percentage of students who met the standard was only 72%. Unfortunately, this percentage dropped to 71% in 2012, then 65% in 2013. Even more troubling than the obvious decline is that the percentage of students reported by the Department as *approaching* or *meeting* the standard is a composite number that may not very well represent different skill sets. Further, it is unclear what "approaching" the standard really means. Based on the statistics personally collected by email request from the Instructional Development and Achievement Office, in 2011, 72% of students met the expectation, with 11% of students failing to achieve standard but described as *approaching the standard*. Similar situations were found from the score report in 2012 and 2013 as well. In contrast, there has been a dramatic growth in the level of experienced difficulty from 17% in 2011 to 29% in 2013.

When comparing students' mathematics performance between the English and the French School Boards, students from the French school board were scoring higher than their counterparts in the English School Board in terms of the overall average as well as the proportion of students scoring at each level. Table 3.17 shows PE student average scores on the provincial Grade 3 assessment of mathematics through 2010-2014, comparing both the English/French Immersion (EN/FI) and French (CSLF) students. Table 3.18 shows the percentage of students at different levels for the same period and assessments.

Table 3.17

PE Student Average Scores in Grade 3 Mathematics Assessments (Primary Mathematics Assessment) from 2010 to 2014

School	Average Score			
	2010	2011	2012	2013
EN and FI	78	79	77	75
CSLF	81	86	81	82
Provincial	78	80	78	75

Note. Adapted from the Instructional Development & Achievement Office of the Department of Education and Early Childhood Development, PE (2010, 2011, 2012, 2013).

Table 3.18

Percentage of Students in Each Level of Achievement in Grade 3 Mathematics Assessments (Primary Mathematics Assessment) from 2010 to 2014

	2010	2011	2012	2013
English Language School Board				
Met expectations	67	71	71	64
Approached expectations	12	11	5	6
Experienced difficulty	21	18	24	30
French Language School Board				
Met expectations	75	85	80	89
Approached expectations	15	8	2	3
Experienced difficulty	10	7	18	8
Provincial				
Met expectations	68	72	71	65
Approached expectations	12	11	6	6
Experienced difficulty	20	17	23	29

Note. Adapted from provincial assessment reports from the Instructional Development & Achievement Office of the Department of Education and Early Childhood Development, PE (2010, 2011, 2012, 2013).

Grade 6 Mathematics. Grade 6 mathematics assessment is fairly new to the province as it was introduced in 2012. Results from the first Grade 6 mathematics assessment were not good, with only 56% students who met the standard and 5% of students who were approaching the standard, with 39% of the students failed to meet the standard. There was a considerable growth in the number of students who were at or approaching the standard in 2014, which was 75% compared to 61% in the previous year (PE Department of

Education, Early Learning and Culture, 2014).

With respect to student performance between the two School Boards, students from the English School Board had a higher percentage of students meeting the standard and fewer students approaching the standard. However, there was not much difference in student proportions in the lowest level of achievement described as those who experienced difficulty, which was 38% and 39% respectively. Table 3.19 highlights the percentage of students at each level on the Grade 6 mathematics assessment for 2013.

Table 3.19

Percent of Students in each Level of Achievement in the Grade 6 Mathematics Assessment (Elementary Mathematics Assessment) in 2013

School Board	Met Expectations	Approached Expectations	Experienced Difficulty
ELSB	57	5	38
CSLF	51	10	39
Provincial	56	5	39

Note. Adapted from provincial assessment reports from the Instruction Development & Achievement Office of the Department of Education and Early Childhood Development, PE (2013).

Grade 9 mathematics. The same situation happened with PE's LSA of Grade 9 Mathematics (Secondary Mathematics Assessment). After being implemented in 2007, there was an increase in the average score from 59 (2007) to 71 (2013) in English School Board. In the same period, scores from PE students' PCAP mathematics assessments from 2010 to 2013 also increased from 460 to 492, which made PE's national ranking jump from the last jurisdiction (eleventh) to the fourth jurisdiction. However, there were still over one third of the students who were facing difficulties in the assessment according to 2013 provincial assessment report, with 32% of students in both English and French School Board reported as *experiencing difficulty*. Table 3.20 shows the average score of PE students on the Grade 9 Mathematics assessment from 2008 to 2013.

Table 3.20

Average Score in Grade 9 Mathematics Assessments (Intermediate Mathematics Assessment) from 2008 to 2013

School Board	2008	2009	2010	2011	2012	2013
ESD	59	64	64	56	66	71
WSB	59	59	65	56	66	
CSLF	67	56	61	58	65	71
Province	59	62	64	56	66	-

Note. ESD refers to Eastern School Board. WSB refers to Western School Board. CSLF refers to French Program in PE public education. Adapted from provincial assessment reports from the Instruction Development & Achievement Office of the Department of Education, Early Learning, and Culture, PE (2008, 2009, 2010, 2011, 2012 and 2013).

A map of student progress in mathematics can be traced using the provincial assessment results. The map reveals that two cohorts of Grade 3 students who enrolled in 2010 and 2011 experienced a sharp drop in the number of students who met the expectations when they moved to Grade 6 in 2013 and 2014 respectively. On the Grade 3 mathematics assessment in 2010, 68% of the students met the standard, however, only 56% of the students in the same cohort were at the same level when they reached Grade 6 in 2013. For the students enrolled in Grade 3 in 2011, 72% achieved the expected level; when they took the Grade 6 assessment three years later, the 2014 annual report showed that 75% of the students met or approached the standard. At first glance it appears that there was a slight improvement in mathematics competencies in this cohort from Grade 3 to Grade 6. However, when compared with the detailed data from the Instructional Development and Achievement Office of the Department of Education, the actual percentage of students who met the standard in Grade 6 mathematics assessment was 70.5%, with 4.8% of the students approaching the standard (personal communication with B. Baily, November 18, 2015). Therefore, it is reasonable to say that there was no improvement in student achievement in

mathematics, instead, the percentage of students who met the standard dropped slightly from Grade 3 to Grade 6.

The mapping of student achievement in mathematics is hindered in two ways. Firstly, as noted before, PE was the last province to initiate the provincial assessments and they have not been implemented systematically. Students' performance on provincial assessments can be compared with the data released in the annual reports of different cohorts of students in the same grade for the same year. However, it is not possible to compare the same cohort of students from Grades 3, 6, and 9 to map their progress in mathematics to provide an indicator of whether students have improved or not. Based on the analysis discussed above, only two cohorts of students' assessment scores can be traced back for comparison, the Grade 3 cohort of 2010 and that of the 2011 cohort. It is hard to see any progress in students' knowledge and skills in mathematics from the test results. In fact, there was a dramatic drop of students (12%) between those who were at the standard in the first cohort of Grade 3 students and those who met the standard on the Grade 6 mathematic assessment. Figure 3.4 demonstrates the achievement of the two cohorts of students in mathematics from Grade 3 to 6 for the English program.

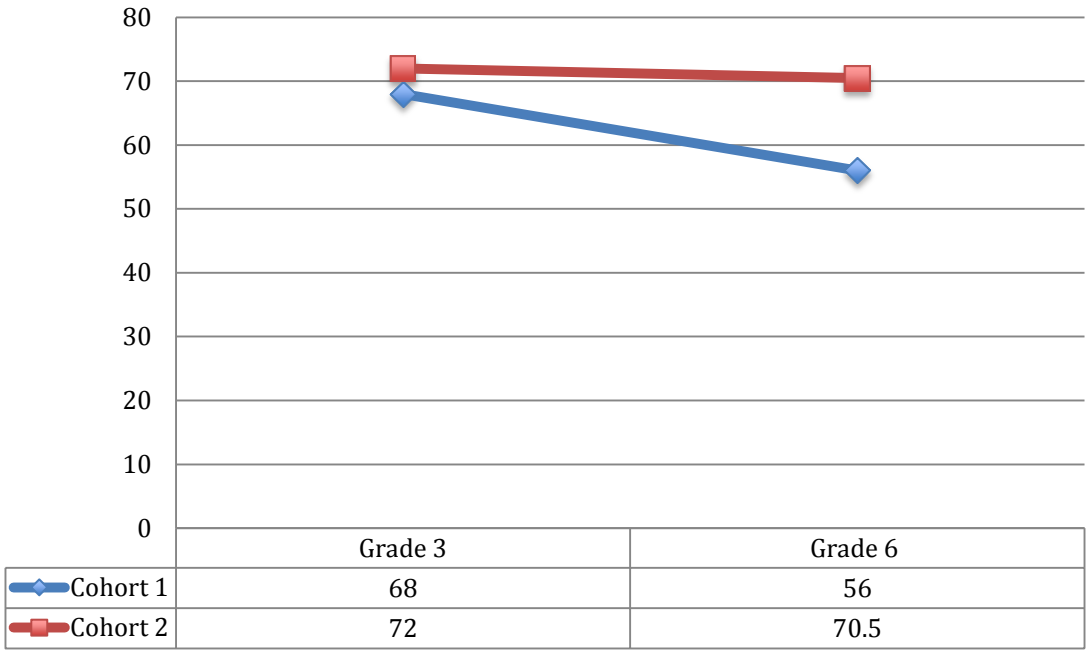


Figure 3.4. Percent of Students who Met the Standard in Mathmatics on Provincial Assessment in Two Cohorts of Students. Adapted from provincial assessment reports from the Instruction Development & Achievment Office of Department of Education, Early learning, and Culture (2010, 2011, 2013, 2014).

Secondly, the inconsistency of data from the reports obtained from the Department of Education's website and the Instructional Development and Achievement office of the Department of Education make it quite difficult to examine students' achievement over the period. Since the government website contained a composite score representing the percentage of students who were both at or approaching the standard, it is difficult for the public to obtain a clear understanding of the exact percentage of students who met the expected level on the provincial scale and those who did not. Moreover, some data cannot be obtained from the online reports since they have not been released to the public.

Fortunately, in the data gathering phase of the study, I was able to obtain additional data from the Instructional Development and Achievement Office of the Department of Education who were able to provide the exact percentage of students who passed the required

level and how many of the students were approaching it but this is not public data. The difficulty with accessing PE LSA data was further complicated given the varying method of reporting student achievement measure. For example, the detailed data contained percentages of students who met the standard, approached to the standard, and had difficulties for Grade 3 students from 2010 to 2013, Grade 6 students from 2013 to 2014, and Grade 9 students in 2013. However, the Grade 9 assessment scores were presented as average scores from 2008 to 2012, without any further explanation about the percentage of students scoring in each achievement level.

It is important to note that even with accurate data representing the percent of students who met the standards on each of the provincial LSAs, the standard of each assessment is not explained with any information, which made it rather difficult to describe the knowledge and the skills the students are expected to achieve and how this provincial standard aligned with the PCAP and PISA levels; which still remains unknown. Given that Grade 9 provincial averages were lower than the Grade 6 and 3 overall averages, one might wonder whether the publication of student scores by achievement levels was concealed because the scores were too low. Regardless of why the Grade 9 scores were not presented by achievement levels, the issue at hand is the lack of transparency and accountability of PE's LSA body.

As previously noted, it is challenging to align the PE assessment standards with the national and international LSAs or map cohorts of students' achievement from the provincial to the national and international LSAs. One of the main reasons was that PE started its provincial assessments in the senior grades rather late compared to their early participation in the PCAP and PISA. The only related data is from the 2013 PCAP assessment. The cohort of students who wrote the Grade 3 provincial assessments in 2008 were randomly

selected to write the 2013 PCAP assessment (i.e., in Grade 8). When examining this cohort of students, there was a drop in the percentage of students who met the standard in reading when they wrote the assessment in Grade 6 in 2011. This group of students wrote the PCAP for reading in Grade 8 in 2013, where they ranked sixth out of the ten provinces. This pattern was echoed in the Grade 9 reading scores, which showed that progress was made where the percentage of students who met the standard increased to 81% from 66%. This is the only cohort of students that can be mapped to align student achievement between the provincial and national assessments. In terms of the low ranking that PE students scored in PISA, no provincial data can be obtained to explain the situation due to the absence of a provincial assessment in the senior grade. When considering the existing data, it appears that there is a high percentage of students who met the standards. However, it should be noted that without any descriptions and explanations of the standards in provincial assessments, it is not possible for the public or the stakeholders know exactly what knowledge and skills that students are excelling in. Furthermore, given the absence of transparency in identifying the provincial standards for each LSA, it is unknown whether the provincial standards are aligned with classroom standards or standards on the national and international LSAs. Hence, it is difficult to infer any connections in student achievement in these three types of LSAs which affects the accountability function of LSAs.

Conclusion

This section reviewed the history and current practices of LSAs on PE. Based on the three types of LSAs' results, PE students did not make large gains in mathematics when comparing achievement from year to year or when comparing cohorts of students through Grades 3, 6, and 9. There has been notable progress in reading for Grade 3 students.

However, the number of students who met the expectation in reading did not increase through the years, instead, it decreased considerably when the same cohort reached Grade 6. It is important to note that there has not been any provincial assessment in science for PE students, so it is not possible to link provincial performance in science with PISA and PCAP.

Student achievement in reading and mathematics was mapped using two types of LSA data drawn from the annual reports presented on the PE Department of Education website and from personal requests to the Instructional Development and Achievement Office of the Department of Education. Consistent trends can be seen when comparing data from year-to-year at key grades. Given the high percentage of students who met the standard as released in the annual reports, a false impression is given to stakeholders such as parents, teachers, and principals. More importantly, it remains unclear what the provincial standards are for each assessments. Without any detailed information explaining the provincial standard, there are no criteria to describe students who met the standard. Although PE has a provincial LSA program which can fulfill the accountability role of LSAs but without accurate and transparent information related to student performance or standards, the LSA program fails to hold those responsible for educating the public accountable.

CHAPTER FOUR

Methods

The purpose of this study was to examine public knowledge about LSAs and the perceptions they hold for these instruments. Participants' knowledge and perceptions of LSAs were based on Nagy's (2000) framework for LSAs, that posits three functions for LSA: accountability, gatekeeping, and instructional improvement. In addition, a fourth function, monitoring student achievement, is another function of LSAs stated in LSA programs in Canada (Klinger, DeLuca, & Miller, 2008). The information gathered in this study acts as indicators of the functionality of LSAs in the province of PE and by extension, the utility of LSA data as indicators of the wellbeing of the education system and the province's economic prosperity. A questionnaire (see Appendix A) was used to gather information about LSA knowledge and perceptions in these areas. Four follow-up interviews were also conducted to extend the findings from the questionnaire. Questionnaires and interviews are often used in combination to explore issues in education (e.g., Brookhart & Durkin, 2003; Miller, 2013). However, it is important to ensure the construct is well aligned in both instruments (Brown & Harris, 2010). The following sections describe the design of the research instrument, the participants involved in the study, and the methods for data analysis. The methods for analysis are organized according to descriptive and inferential statistical analysis. The different types of inferential analysis used included factor analysis and analysis of variance (ANOVA). The method for analyzing interview data concludes this chapter.

Conceptualization of Research

It is important to acknowledge that this study was conceptualized in my Master of Education quantitative statistics course. This course adopted a project-based approach to exploring quantitative statistics where each of the five students were required to select an area of study under the umbrella of educational culture. For me, this area of study was public knowledge and perceptions related to LSAs. As a member of this broader research team, I was required to contribute to the project proposal and ethics application. The conception of the research focusing on LSA was my own as was the design of the instrument, data analysis, and writing of this thesis.

Research Design

This study employed an explanatory sequential mixed methods (Christensen & Johnson, 2013). The mixed sequential explanatory design consists of two distinct phases: quantitative followed by qualitative (Creswell, 2003). In the quantitative phase, researchers collect and analyze the data, followed by collecting and analyzing the qualitative data in the second phase to help explain, support, or elaborate on the themes raised in the quantitative portion. The quantitative data provide a general understanding of the research question while the qualitative data refine and explain the statistical results by exploring participants' views in more depth (Creswell, 2003; Rossman & Wilson, 1985; Teddlie & Tashakkori, 2003). The first stage of research involved designing a construct for the beliefs and opinions being studied, followed by creating an instrument to gather data based on the construct. An electronic questionnaire was selected over a paper and pencil questionnaire given the advantage of the internet in reaching groups of individuals who are difficult to approach combined with the ease to implement, quick, and economic characteristics of the

electronic questionnaire (Garton, Haythornthwaite, & Wellman, 1999; Wellman, 1997). Paper and pencil questionnaires were also distributed as a back-up strategy to obtain a large enough sample size during the final two weeks of the online survey.

Next, a strategy for soliciting participants to complete the electronic questionnaire was formalized as a proactive means to solicit a large sample size to represent the views of PE residents. Following the data collection, the data were analyzed. To broaden the quantitative data and extend the findings, in-depth perceptions on LSAs were gathered by interviewing four key educational stakeholders representing the provincial government, high schools, the Home and School Federation, and the economic sector.

Questionnaire Development

The questionnaire contained three major sections: demographics (10 items – surveying 6 areas of inquiry), LSA knowledge (6 items), and perceptions about importance/influence of LSA (17 items).

Construct

The construct focused on participants' perceptions towards LSAs. Perceptions are shaped by ones cultural environment and it allows people to make sense of the events such as LSAs in their environment. Perceptions are important because behaviour is based on perceptions of reality or what people consider normal or good. Perceptions are influenced by a number of factors such as interests, experiences, or proximity. When exploring perceptions towards LSAs, it is likely that they are influenced by their proximity to the education system. Hence, teachers and parents would be quite proximal and would have perceptions shaped by factors different from non-parents and may hold a knowledge and experience set different from others. Perceptions of people in businesses may also be more

informed about LSAs if businesses understand the relationship between education and economic prosperity.

Factors influencing participants' perceptions towards LSAs were organized by importance of LSAs in providing different types of information and the relevance of LSA information. The context for questionnaire items exploring participants' perceptions towards LSAs focused on students, teachers, education system, businesses, and other people in general. In total there were 12 items exploring the construct with the following focus: (2 items) education system, (4 items) students, (2 items) businesses, (2 items) teachers, (2 items) other people.

Demographic Items

Participants were asked to identify their age and education. In Canada, 21% of the baby boomers hold a university certificate, and a further 16% of them hold a high school diploma (Community and Health Services, 2014). It is notable that baby boomer women are more educated in Canada, where 16% of the second-wave baby boomer women hold college degrees compared to 5% of the pre-baby boomers who held only post-secondary education credentials (Galarneau, 1994). Participants' age, gender, and employment status, and occupation was collected to describe the sample and compare the sample with the greater PE population. Participants' occupations were documented based on Statistics Canada's National Occupation Codes. The classification structure has 10 different kinds of occupations listed as: (0) management occupation; (1) business, finance and administration occupations; (2) Natural and applied sciences and related occupations; (3) Health occupations; (4) Occupations in education, law and social, community and government services; (5) Occupations in art, culture, recreation and sport; (6) Sales and service

occupations; (7) Trades, transport and equipment operators and related occupations; (8) Natural resources, agriculture and related production occupations; and (9) Occupations in manufacturing and utilities (Canada Statistics, 2011a).

In addition, participants were asked to indicate their parental status, educational background, and cultural affiliations group. These three variables also described participants' demographic characteristics but were also used as the three independent variables. Information about parental status was solicited because according to the PISA report, parental status was an important indicator that may influence student achievement in reading (CMEC, 2001). Participants with a child or children currently in the education system were referred to as parents and participants who did not have any children or currently did not have any children in the education system were referred to as non-parents in that they were less likely to be as informed about children's' LSA. It is possible that in this coding system parents whose child may have recently graduated from the kindergarten to Grade 12 system would be excluded from the data set; however, it was not feasible to set another exclusion point since there is no ideal stopping point.

As previously noted, participants' educational background was gathered because it was assumed that there would be differences between responses from participants with different levels of education, namely, participants with higher education credentials were likely to be more informed about LSAs and thus have more positive perceptions towards LSAs. Participants were asked to identify their educational background as belonging to the following six groups: below grade 12, grade 12, college/apprenticeship, undergraduate, master, and PhD.

Lastly, in PE there are four distinct cultural groups: French, Aboriginal,

Migrants/Immigrants (Newcomers)[9], and people born in PE, commonly described as Islanders. Participants were asked to identify their cultural affiliation because as documented in the literature review, there was large gap in students' LSA scores due to immigrant status. In addition to that, there were concerns about the challenges that English language learners faced on LSAs even if they receive adaptations. Based on the findings presented in the literature review, it is quite possible that participants' cultural affiliations would influence their perceptions about LSAs. It was anticipated that the Newcomers' perceptions on LSAs would be more positive than people from the local community.

To summarize, it was hypothesized that if participants had children in the education system, they would be more apt to be informed about students' achievements on LSAs in comparison to participants who do not have children in the education system. Participants holding higher education credentials were also thought to be more informed about LSAs in comparison to those who held lesser or no credentials. This rationale extends from the CMEC's research, which showed that parents who have higher educational credentials had positive influences on children's performance in LSAs (CMEC, 2004). Thus it was believed that these participants were more likely to be informed of the trends and issues in education relating to LSAs. The last independent variable was participants' cultural affiliation. As found by the CMEC (2007), there was significant difference in LSA results between immigrant and non-immigrant students in science performance in Quebec, Manitoba and British Columbia. Differences between other cultural groups in PE (i.e., Francophone, Aboriginal) are unknown as there was no literature examining different perceptions on LSAs based on ones cultural affiliation; however, given that these communities are culturally

[9] The term *Newcomer* is a common expression in PE to refer to immigrants or migrants new to PE.

diverse, it is reasonable that their cultural differences may also influence their perceptions about LSA.

LSA Knowledge Items

Items surveying participants' knowledge about LSAs were structured based on the PISA's four domains of assessment (reading, math, science, and problem-solving). Participants were asked to identify how PE students scored in relation to students from other participating jurisdictions in these four areas. Another two items were presented to ask the participants to identify how PE student achievement compared to the average performance of other provinces on national and international LSAs. Information from these six items would provide an indicator of participants' knowledge of PE student achievement.

The perceptions of LSAs. The first three items surveyed participants regarding their perceived significance of LSAs at the provincial, national, and international levels in providing information about PE student achievement. Next, the items examining the construct were presented. The first set of six items explored whether participants believed that LSAs provided information about: (a) the education system, (b) students aptitude for success at college or university, (c) students ability to solve problems in the world around them, (d) students' skills, (e) successfulness of businesses on PE, and (f) whether teachers were doing their jobs. Participants responded to these six items using a six-point rating scale anchored at each end with "(1) Tells us a lot" through to "(6) Does not say anything". This even-numbered scale was designed to force participants to express their views on the subject matter, since most people are legitimately neutral on a subject, which is not considered valuable to investigators (Christensen & Johnson, 2013).

Following this set of items was another set of six items that were part of the construct that focused on perceptions related to information provided in the PISA reports. These items focused on information related to children's success and the quality of schools and teachers. An additional item garnered perceptions about whether businesses thinking of moving to Canada examined the PISA reports. These items also applied a six-point rating scale anchored at one end with "(1)strongly agree" and "(6)strongly disagree" at the other end of the scale. Lastly, a dichotomous item explored whether participants would move their child to a school that had better provincial test scores if they could. This item was not part of the scale but rather probed the extent to which parents were affected by LSA results in their school community.

In total, there were 10 demographic items of which three were used as grouping variables (independent variables). There was an additional six items that explored participants general knowledge about PE student achievement on LSAs. The last part of the questionnaire contained three items to survey participants' perceptions of three types of LSAs followed by a set of 12 items (i.e., two groups of six items) designed to explore the construct.

The questionnaire was piloted over a two-week period where friends, family members, and academic peers were asked to respond to the questionnaire. Minor changes to wording were made to enhance the viability to comprehend the items before distribution to the PE public.

Sampling Design

The total population of Prince Edward Island is 146, 283 (Government of Prince Edward Island, 2014a). With a confidence interval of 95%, and a 5% margin of error, the

sample necessary for statistical accuracy was determined to be 384 participants. To solicit the required number of participants and target specific groups within the community, five organizations were contacted to facilitate the distribution of the questionnaire. These groups included the: PEI Association for Newcomers to Canada, Employment Journey newspaper, Prince Edward Island Retired Teachers' Association, Education 20/20, and Prince Island Chamber of Commerce. Unfortunately, the Prince Edward Island Teachers Federation did not respond to the request to facilitate the distribution of the questionnaire. Since the perceptions of teachers was considered valuable in this study, the Retired Teachers Association were subsequently contacted and agreed to distribute the questionnaire to their membership. Additionally, a link to the online questionnaire was posted on the websites of the local home and school associations and the University of Prince Edward, which allowed students, parents, professors, and communities to participate in the study. The online questionnaire was also sent to personal friends, colleagues, neighbours, with the request that they send the questionnaire to their personal contacts and so forth.

The last strategy for boosting the sample size was resorting to using paper and pencil questionnaires. Questionnaires were distributed in a small shopping mall (e.g., Charlottetown Mall), open market (e.g. Farmers Market), community event sites (e.g., events at the Confederation Centre) as well as restaurants, walk-in clinics, coffee shops, and the college campus.

Questionnaire Distribution Methods

In the past, online questionnaires were thought to only reach middle and upper classes that could afford internet related charges. Since the early 2000s increasingly more people have access to the internet, which has attracted researchers to online questionnaires given the

greater accessibility to the population (Gosling, Vazire, Srivastava, & John, 2004). Online questionnaires are even more attractive given the number of list serves and news or community groups that organizations use to communicate to their members (Kraut, Olson, Banaji, Brukman, Cohen, & Couper, 2004). This greater access to the population combined with easier, quicker, and cheaper aspects of online questionnaires made electronic questionnaires the most attractive method for distributing this questionnaire (Skarupova, 2014).

The electronic questionnaire was distributed to five key organizations. The first group was PEI Association for Newcomers, which has a list serve of approximately 2000 members. Second was the *Employment Journey* newspaper, which freely distribute their newspaper to over 30, 000 residents of PE. A full-page article describing this study was featured along with a link to the electronic questionnaire. The link to the questionnaire was also distributed to the Retired Teachers' Association, which has a membership of approximately 1000 members, and to an educational advocacy group known as Education 20/20, which has approximately 100 members. Lastly, the University of Prince Edward Island posted a headliner story about this research project and solicited members of the university community to complete the questionnaire.

At week two into the data collection period, the response rate was not as prolific as desired and a decision was made to distribute paper questionnaires in places where people would be most likely to complete them. Given that it was important to hear from all people in PE, a creative approach was adopted in identifying places where people spent time waiting as it was believed these people would be more apt to complete the questionnaire as they waited: hair salons, medical offices, restaurants, public activity centers, and the book market.

This initiative resulted in an additional 93 paper questionnaires. Combining the paper questionnaires with a final tally of 422 online questionnaires, the sample size of this study was 515 participants, well over the required minimum.

While questionnaires, electronic or paper and pencil, were appealing, it was important to be cognizant of the short-comings of this data collection strategy. Given that people who completed the questionnaires were self-selected, the issue of representation was raised (Matsuo, McIntyre, Tomazic, & Katz, 2004). To examine the representativeness of our sample, the demographic characteristics of the sample was compared to the demographic characteristics known about the PE population (e.g., gender, age, employment status). Table 4.1[10] below compares seven demographic characteristics between the sample and population (Government of PE, 2013; Statistics Canada, 2015a). Although it is near impossible to duplicate the representativeness of a population in a sample, the sample used in this study was sufficiently large and representative of the key cultural groups and characteristics examined, except that in the sample, the Newcomers group and participants who had higher education attainment such as undergraduate and graduate degrees were over represented. Hence caution was exercised in the analysis that examined patterns of responses for these groups.

[10] Data was adapted from 40th Statistics Review of Prince Edward Island 2013, released by June, 2014. The population by then was 145,237. Data with * were drawn from Statistics Canada National Household Survey, 2011.

Table 4.1

Demographic Statistics of the Participants Compared to PE Population

	Sample	General PE Population
Age		
Born before 1945	14 (2.7)	16,297 (11.2)
Born from 1945 to 1964	144 (28.2)	41,699(28.7)
Born after 1964	353 (69.1)	87,241 (60.0)
Gender		
Male	128 (25.4)	70,695 (48.6)
Female	376 (74.6)	74,542 (51.3)
Education attainment (>15-year)*		
Below Grade 12	3 (0.6)	24,855 (21.7)
Grade 12	51 (10.8)	29,970 (26.2)
College/Apprenticeship	111 (23.5)	59,375 (51.9)
Undergraduate	145 (30.7)	13,195 (11.5)
Graduate (masters and PhD)	162 (34.3)	6,250 (5.4)
Employment rate*	275 (55.9)	68,635 (59.8)
French*	37 (7.7)	28,950 (20.6)
Aboriginal*	18 (3.7)	4,460 (3.1)
Immigrants*	121 (25.1)	7,090 (5.0)

Questionnaire Analysis

Data from the electronic questionnaire was transferred directly to SPSS (Statistical Packages for the Social Science), which avoided the potential for data entry errors. The paper and pencil questionnaires were then added to this data set. The explore function in SPSS was used to ensure that data entered manually was in the appropriate ranges, thus minimizing the presence of data entry errors.

Data Analysis

Descriptive Statistics

Descriptive analysis included calculating frequency and percentages for each item. For ordinal items, the mean and standard deviation were also calculated. A report of the descriptive analysis is presented in Appendix C.

To examine patterns between two or more categorical variables, a Chi-square test for independence was performed which compares the observed frequencies of participants' responses that occur in each of the categories. The Chi-square test generates a cross-tabulation table showing this distribution. To use the Chi-square test, the minimum expected cell frequency must be 5 or greater. The Pearson chi-square value must be less than 0.05 to be significant. The effect size statistic for a two by two cross-tabulation is the phi coefficient which ranges from 0 to 1 where 0.10 is a small effect, 0.30 is a medium effect, and 0.50 is a large effect (Cohen, 1988). If a larger table was used (e.g., 2 x 3), Cramer's V would be reported where 0.07 is a small effect, 0.21 is a medium effect and 0.35 is a large effect. If the cross-tabulation was larger (e.g., 3 x 4 table), then 0.06 is a small effect, 0.17 is a medium effect, and 0.29 is a large effect (Gravetter & Wallnau, 2004).

Inferential Statistics

Factor analysis. Two types of inferential statistics were used in this study, factor analysis and an analysis of variance (ANOVA). Factor analysis is believed to be an effective statistical tool to reduce the observational and measurable variables to fewer latent variables that share a common variance, which is known as the "reduction of dimension" (Bartholomew, Knott, & Moustaki, 2011). Factor analysis adopts mathematical procedures to simplify the interrelated measures to discover the patterns in a set of variables which is the aim of this method (Childs, 2006; Harman, 1976). The two main types of factor analysis are exploratory factor analysis and confirmatory factor analysis. Exploratory factor analysis was selected for this study because it allows researchers to explore the main dimension or model from large set of latent constructs often represent by a set of items.

The two primary conditions underpinning the use of factor analysis are the sample size and the strength of the relationship between items. Psychometricians vary on the minimum sample size required for factor analysis (Stevens, 1996; Tabachnick & Fidell, 2013), while others argue that the ratio of participants to items is of greater concern (Nunnally, 1978; Tabachnick & Fidell, 2012). Given that the sample size of 515 exceeded the minimum 150 and the ratio of participants to items is 43 to 1, exceeding the recommended 10 to 1 ratio (Nunnally, 1978), the sample used in this thesis well exceeded the minimum criteria. The second condition for using factor analysis requires that the items be inter-correlated with correlation coefficients greater than 0.3. An additional two statistical measures are generated by SPSS to confirm the suitability of the data for factor analysis. The first measure is Bartlett's test of sphericity, which must be significant ($p<0.05$) and the Kaiser-Meyer-Olkin measure must exceed 0.6. Conditions for each of these statistical measures were met and reported in the following chapter.

ANOVA. A three-way analysis of variance (ANOVA) was used to examine the effect on the three variables on the independent variable (perceptions towards LSAs). ANOVA generates a value known as the *F* ratio. This value is calculated by dividing the variance between the two groups by the variance within the groups. A large *F* ratio signals more variability between the groups, which is due to the independent variable. A significant *F* ratio ($p<0.05$) means the null hypothesis is rejected, thereby acknowledging that the groups are significantly different. The first step when interpreting the output of a three-way ANOVA is to determine whether there is any interaction effects between the independent variables. Next, the main effects are examined which compares the relationship between the three independent variables and the one dependent variables. In

the event there is a significant difference (p<0.05) and there are more than two groups (e.g., age has three groups/categories), a post hoc analysis is performed to determine which groups significantly differ (DeCoster, 2006; Kao & Green, 2008;).

There are a number of different post hoc analysis. The most common are Tukey's Honestly Significant Difference and the Scheffe test (both named after psychometricians). Since the Sheffe test has a tendency to be too cautious (Klockars & Sax, 1986), Tukey's Honestly Significant Difference was selected to determine which groups differed in this study.

Following a significant difference, the effect size must also be calculated because large samples can generate significant differences that do not necessarily have any practical or theoretical differences (Tabachnick & Fidell, 2012). This analysis examines how large or small the difference between the mean scores. Although there are a number of different ways to calculate effect size, eta squared, $\eta = {SS_{between}}/{SS_{total}}$, was used in this study.

Cohen's (1988) guide to interpreting the strength of the effect size was followed where an effect size of 0.01 was small, 0.06 was medium, and 0.14 was large.

Before using the ANOVA, there are six assumptions underpinning its use. The first assumption specifies that the level of parametric measurement be at the interval or ratio level on a continuous scale. In this thesis, a six-point continuous rating scale was used for all 12 items. The second assumption assumes that the sample is random, which is seldom the case in social science research and ANOVA is considered robust enough to handle a non-random sample. ANOVA's regulations state that the independent variables should consist of two or more categorically independent groups. In this instance, the independent variables that had more than two independent groups included age (e.g., 3 groups: pre-baby boomer, 1945-

1965, and post-baby boomer) and level of education (e.g., 6 groups: below grade 12, grade12, college, undergraduates, graduates (Master), and graduates (PhD), and so forth. The third assumption emphasizes the independence of observations in each group or between the groups. For example, there must be different participants in each group with no participants being in more than one group. The rules also regulate that there should not be any outliers, which is not a problem for electronic questionnaires. Paper pencil questionnaires should be checked for data errors in the manual entering process. ANOVA also requires normal distribution for each category of the independent variables. This can be easily examined by using the explore function in SPSS. The last assumption requires a range of homogeneity variances, which calls for the variances or the responses to be homogenous for each of the groups. Levene's test was conducted to examine this factor, showing the statistical result to be greater than 0.05.

Interviews

To elaborate on the questionnaire findings, in-depth interviews were conducted with key people related to their special areas and interests in LSA. These interviews also provided opportunities to gather more information on the growth of LSAs or changes to LSA practice in PE. In-depth interviews are considered to be useful in gathering detailed information when investigating new issues or to further explore specific areas, while being flexible enough to uncover new areas, in this case regarding LSA in PE (Boyce & Neale, 2006; Legard, Keegan & Ward, 2003).

Participants for each of these areas were selected based on their expertise and leadership within the area as well as availability to participate in an interview. Participants were initially contacted by email with a letter of information, which solicited their

participation in the study. The four participants selected for interviews represented the following organizations: (a) PE Department of Education (LSA specialist), (b) high schools, (c) Home and School Federation, and (d) Bioscience Business Sector. After receiving the comfirmation about their willingness to participate in the study, the written consent form and the interview protocols with background information were presented to the participants. After with participants' signatures on the consent form, the interviews were conducted.

Interview Participants

Participant 1 was an Achievement Specialist from the Instructional Development and Achievement Office of the Department of Education and Early Childhood Development of PE. This division of PE's Department of Education and Early Childhood Development focuses on developing and contributing to teachers' professional development and is responsible for administering and scoring the provincial LSAs referred to as the *Common Assessments* in PE. The achievement specialist also provides leadership, guidance, and support to the Department, principals, and teachers in the area of student assessment. One of the key responsibilities of this participant was to lead teachers in the development and administration of the provincial LSAs while overseeing national and international student assessment programs and activities (Department of Education, Early Learning and Culture, 2015g).

Participant 2 was a high school principal in the western part of PE, who had been working as a principal for four years. Prior to this position, the participant spent one year as a vice-principal as well as several years as a Grade 11 resource teacher. Participant 2 had recently participated in the piloting of Grade 10 Secondary Literacy Assessment and Grade

11 Mathematics assessments. As a school principal, this participant annually examined incoming students' Grade 9 assessment of mathematics data.

Participant 3 was a representative from PEI Home and School Federation, as well as a member of the Conference Board of Canada and professor at the University of Prince Edward Island. This participant was asked to be interviewed as a representative of the Home and School Federation; however, it is likely that his other roles also influenced his responses to the interview questions.

The last participant was selected based on his leadership in the PE business sector. This participant represents both PE Biosciences as well as Nautilus Bioscience Canada Inc. He had a 22-year career as a research scientist in product development and director in a multi-national company based in Singapore. At the time of interview, he chaired the vice president Business Development of Nautilus Bioscience Canada and supported the bioscience cluster on PE.

Interview Questions

Semi-structured questions were developed after examining the related literature and questionnaire findings. General questions were prepared for all participants related to provincial assessments, then for each interviewee, different detailed questions were posed related to their personal area of expertise (see Appendix A for a copy of the interview questions). The general questions posed for all participants were:

1) To what extent are you aware of the most recent PE student performance in provincial assessment/national assessment/international assessment?
2) Are you aware of the applications or uses of provincial assessment (common assessment), national assessment (PCAP) and international assessment (PISA) in PE?

3) To what extent do you understand the purpose of using LSA as an indicator of the functionality of the education system?

Specific questions were also posed to each participant exploring their connection to the education system, field of expertise, and perceptions towards LSAs in PE. The following topics of discussion were raised with the Achievement Specialist: (a) the alignment of classroom and provincial assessments, (b) provincial assessment standards, c) guidelines for accommodation students with special needs; (d) technology to be incorporated in the forthcoming grade 10 literacy and grade 11 mathematics assessments, (e) financial investment of LSAs, (f) functionality of Grade 10 literacy and Grade 11 mathematics assessments, (g) plans to ensure students receive the basic skills advocated by PISA, and (h) teacher professional development related to LSAs.

Similar topics were raised with the high school principal about the alignment of classroom and provincial assessments and teacher professional development. Given the high-stakes nature of Grade 10 literacy and Grade 11 mathematics assessments, detailed questions about the piloting of these two assessments were posed to the principal who was involved with the piloting. Other questions for the principal were focused on the utilization of LSA data, and the effects and challenges related to incorporating LSA findings to influence teaching practice and school leadership. The interview concluded with the principal's perceptions about improvements in the provincial LSA program.

Questions for the representative from the Home and School Federation and the representative from the Bioscience Business Sector were quite similar. Discussion topics focused on: (a) identifying the skills required in the future workforce, (b) how LSA information was used by general businesses and bioscience business, (c) the cause of poor

LSA student achievement in PE, and lastly, (d) suggestions about how to improve student achievement as well as how to create future leaders in PE. Further questions about the implications of poor student achievement on general business as well as bioscience businesses were also posed to the representative of bioscience business sector.

Public school teachers were considered key stakeholders and a fifth interview was planned. However, two teachers that were identified by a school principal as school leaders in LSA did not respond to four interview requests to participate in this study. Thus the data presented in this component of the thesis is not reflective of teachers' LSA perceptions.

Interview Method

The interviews were audio-recorded using an app called Explain Everything, which allowed the researcher to document field notes that were simultaneously recorded in parallel with the audio recording. Each interview lasted about 45 minutes to one hour in length. Interviews were conducted in a convenient place for the interviewee, such as their office or workplace.

Interview Analysis

Open coding (Holton, 2009) was used to determine the presence of common words or phrases in the transcriptions such that inferences could be made to reflect the views of participants. During a second reading of the transcriptions, a more detailed axial coding (Holton, 2009) was used to confirm pre-determined concepts and categories (e.g., aware of PE's LSA results/unaware; positive view of education/negative/no view). Interview protocol was developed by this semi-structured one-on-one interview to ensure that valuable information was obtained during the process of data collection. Interview themes were developed using the software of Nvivo, which can align the transcript and the fields notes

together to facilitate the analysis. After participants' audio recordings and transcript were uploaded into Nvivo, a text search query helped to trace the interconnections of themes from the entire data set. Reflections were made after all the node or themes were established in the coding part of Nvivo.

CHAPTER FIVE

Findings

This chapter presents the findings from the questionnaire and interviews analysis. The questionnaire findings are organized according to descriptive and inferential statistical analysis. The descriptive statistics shows the frequency and percent of participants' demographic characteristics followed by the frequency, percent, mean, and standard deviation for all interval items. Inferential statistics was used to explore relationships between variables. This includes a report on the factor analysis conducted to explore the dimensionality of the scale used to measure participants' perceptions towards LSAs. Cross-tabulations are also used to compare the frequency distribution between two or more variables producing the chi-square test of significance or Cramer's V (if there were more than two categories). In addition, findings from the analysis of variance report on the differences between mean scores. Lastly, common themes found when analyzing the interview data are also summarized.

Questionnaire Findings

Questionnaire demographics. Participants were asked to identify their demographic characteristics in terms of: (a) age, (b) gender, (c) parental status, (d) employment status, (e) occupation, (f) educational background, and (g) cultural affiliation. Five hundred and eleven participants indicated their ages in the three groups. The fewest number of participants was pre-baby boomers, who were born before 1945 (14, 2.7%). Next, 144 (28.2%) were in the middle age group, which is between 51 and 70 years of age. The youngest category who were born after 1964 had the majority of participants (353, 69.1%). The distribution of participants in this study compares to the current age

distribution on PE showing the distribution of ages as 4.2%, 29.0%, and 43.5%, respectively, in the three age groups surveyed by Statistics Canada (2015a).

The majority of the 504 participants indicated their gender was female (female: 376 [73%]; male: 128 [24.9%]). Although this female to male ratio is not reflective of the PE population it is not uncommon in social science research, where females have a tendency to participate more often than males (Curtin, Presser, & Singer 2000; More & Tarnai, 2002; Singer, van Hoewyk, & Maher, 2000). Next, participants were asked to indicate their employment status. Responses to this item were simplified to employed, unemployed, and retired to allow for comparison of this sample with the population. Of these three groups, 399 (81.1%) were employed, 63 (10.8%) unemployed, and 30 (6.1%) were retired. For participants who were employed, 275 (55.9%) had full time jobs, 53 (10.8%) had part-time jobs, and 65 (13.2%) were self-employed. This distribution of participants based on their employment status is similar to the PE population but contains a slightly higher representation of the employed people, where as in the population of PE, only 61.4% are employed (Statistics Canada, 2015c).

Participants were also asked to identify their occupations based on the occupational classification chart from Statistics Canada (2015b). Of the 474 participants who provided a response, almost half (235, 49.6%) indicated they were working in education, law, or social services fields. Another 66 (13.9%) participants were working in the field of business, finance, or administration, and 48 (10.1%) participants were in the field of sales or the service sector. Additionally, 30 (6.3%) participants indicated they were homemakers or stay-home mothers. There was a small percentage of participants in other categories such as natural and applied sciences (21, 4.4%), health (29, 6.1%), art, culture, recreation and sport

(22, 4.6%), trades, transport and equipment operators (8, 1.7%), natural resources, agriculture (12, 2.5%), and manufacturing and utilities (3, 0.6%). According to PE profile from the National Household Survey in 2011 (Statistics Canada, 2015b) that also solicited information based on the occupational chart, among the total population of 78,060 aged 15 and above, in the 2011 labour force, 8,810 (11.2%) worked in the field related to education, law, or social services fields, 11,640 (14.9%) worked in the field of business, finance, or administration and 17,510 (22.4%) were in the field of sales or service sector. There were 3,675 (4.7%) people working in the field of natural and applied sciences, 4,990 (6.3%) in health related fields, 1,620 (2.0%) people worked in art, culture, recreation and sport, 11,950 (15.3%) people were employed in trades, transport, and equipment operation, 6,140 (7.8%) people served in natural resources, agriculture field, and 3,360 (4.3%) in manufacturing and utilities industry. However, there was no data for homemakers or stay-home mothers in the report. Except for the large divergence in the percentage of occupation related to education, law, and social services in the sample (49.6%) compared with PE population (11.2%), the percentage of other occupations were quite similar. Therefore, it is reasonable to assume that the sample well reflected the greater PE community based on this demographic characteristic. This information is summarized in Table 5.1 below.

Table 5.1

Comparison of Sample and Population Based on Occupation Classification Codes

Occupation	Sample (2015)	National Household Survey (2011)
Education, law, or social service	235 (49.6%)	8,810 (11.2%)
Business, finance or administration	66 (13.9%)	11,640 (14.9%)
Sales or the service sector	48 (10.1%)	17,510 (22.4%)
Natural and applied sciences	21 (4.4%)	4,990 (6.3%)
Health	29 (6.1%)	4,990 (6.3%)
Art, culture, recreation and sport	22 (4.6%)	1,620 (2.0%)
Trades, transport and equipment operation	8 (1.7%)	11,950 (15.3%)
Natural resources, agriculture field	12 (2.5%)	6,140 (7.8%)
Manufacturing and utilities industry	3 (0.6%)	3,360 (4.3%)
Stay at home mothers	30 (6.3%)	n/a

Independent variables. Participants' parental status, educational background, and cultural affiliations were considered to be descriptive as well as independent variables that might influence participants' perceptions about LSAs. In terms of parental status, 256 (50.6%) of the 506 participants indicated they had a child in the public education system. Next, participants were asked to identify their highest level of educational attainment in six categories: below Grade 12, Grade 12, apprenticeship/college, undergraduate, graduate (Master), and graduate (PhD). The distribution in each of these categories was: 3 (0.6%), 51 (10.8%), 111 (23.5%), 145 (30.7%), 133 (28.2%), and 29 (6.1%), respectively. When contrasting the distribution of this sample with the educational background obtained from Statistics Canada's National Household Survey 2011 (Statistics Canada, 2015b), the PE population over 15 years of age was 114, 200 of which 24,855 (21.7%) people had no certificate or diploma, 29,970 (26.2%) people held a high school diploma, and 59,375 (51.9%) people held apprenticeship or college certificates. Another 13,195 (11.5%) people held undergraduate degrees, and 6,250 (5.4%) people held post-secondary degrees above the bachelor level (masters or doctoral degrees). The comparison

of the sample with the PE sample from the National Household Survey is summarized in Table 5.2 below. Given the smaller number of participants in the below Grade 12 category, this category was combined with the Grade 12 to represent participants hold Grade 12 certification or less. Similarly, the small number of participants with PhDs was combined with participants holding Master's degrees to represent participants with graduate level certification.

Table 5.2

Comparison of Sample and Population Based on Educational Attainment

Educational Attainment	Sample (2015)	National Household Survey (2011)
Below grade 12	3 (0.6%)	24,855 (21.7%)
Grade 12	51 (10.8%)	29,970 (26.2%)
Apprenticeship or College Diploma	111 (23.5%)	59,375 (51.9%)
Undergraduate Degree	145 (30.7%)	13,195 (11.5%)
Graduate (Master)	133 (28.2%)	*6,250 (5.4%)
Graduate (PhD)	29 (6.1%)	

*The National Household Survey reported the number of Master's and PhD graduates as one entity. Adapted from the National Household Survey 2011 report.

When comparing the educational attainment data from the sample and the statistics from greater PE community, there was a large difference between the population whose educational attainment was below Grade 12 and those who hold Master's and PhD certification. Namely, sample population in this study underrepresents the population of this first group, with only 0.6% of participants whose educational attainment was below Grade 12, while in the PE population, this group of people constitutes 20.5% of the total population. At the higher educational attainment level, the sample contains a rather large proportion of participants holding Master's or PhD certification (34.3%), while the PE population contains 5.1% of people whose educational attainment is above the bachelor level. Although many attempts were made to reach a broad sample (e.g., free newspapers, public

venues) using both electronic and paper questionnaires, the perceptions of people whose educational attainment was below grade 12 are underrepresented while the perceptions of participants holding higher credentials was over represented possibly because the study was stemmed from a university program. It is difficult to infer how this representation distribution would affect participants' responses. However, it is feasible that participants might be more informed about LSAs given their higher educational attainment.

The last variable was cultural affiliation. Of the 482 participants who provided a response, 291 (60.4%) were not affiliated with any cultural group. A further 37 (7.7%) were affiliated with the French heritage community, 121 (25.1%) self-identified as belonging to the Newcomers community, and 18 (3.7%) belonged to the Aboriginal community. An additional 15 (3.1%) participants indicated they belonged to other cultures. In contrast to the data obtained from Statistics Canada, National Household Survey (2011), 21% were of the French origin, 3.2% were Aboriginal and 5.1% were Newcomers. A comparison of these two samples is summarized in Table 5.3 below.

Table 5.3

Comparison of Sample and Population Based on Cultural Affiliation

Cultural Affiliation	Sample (2015)	National Household Survey (2011)
Local	291 (60.4%)	96,875 (70.5%)
French	37 (7.7%)	28,950 (21.0%)
Newcomers	121 (25.1%)	7,090 (5.1%)
Aboriginal	18 (3.7%)	4,460 (3.2%)
Other	15 (3.1)	

Note. Adapted from the National Household Survey 2011 report.

By comparing our sample with the sample from the National Household Survey (2011), there is a higher proportion of Newcomers in the sample used in this study. Although it is

important to recognize this as a limitation to the study, at the same time, it is seldom possible in social science research to obtain a sample that perfectly matches the distribution of different communities represented in a population.

The goal of the study was to obtain a sample of the PE population that reflected the distribution of people in these seven categories. Although the sample of PE participants represented in this study are not identical in the distribution shown in Statistics Canada documents (National Household Survey, Prince Edward Island, 2011), they are close and the large sample size is sufficient to make reasonable generalizations about the PE community based on responses to the questionnaire.

General Knowledge about PE Student Performance on Large-scale Assessment

Two main items each containing sub-items (total of six items) were created to survey participants' knowledge about student performance on LSAs. Item 10 asked participants to indicate how well PE students scored on LSAs and included an option to indicate they did not know how PE students scored. In hindsight, this item was too vague in that it did not ask participants to focus specifically on a particular LSA. This is problematic because on the PCAP, PE students ranked in the middle whereas on the PISA, PE students ranked at the bottom of the nation. Hence, the data gleaned from Item 10 was the number of participants' who indicated they did not know about students' LSA knowledge. Subsequently, Item 15 was the primary indicator of participants' LSA knowledge. Item 15 contained two sub-items (i.e., 15a and 15b) that also explored participants' knowledge about how PE students compared with students from other provinces on the national PCAP and the international PISA instruments. The ranking of PE students (out of 10 jurisdictions) on the most recent PCAP are: mathematics 4th, science 6th, and reading 6th which we summarize as in the

middle or about average (see Table 3.6 to see rankings for the previous PCAP results) (CMEC, 2014; CMEC, 2013). On the most recent PISA 2012 results, PE students' scored at or very near the bottom in each domain which can be summarized as below the provincial average (see Table 3.1).

When focusing on Item 10, it was found that approximately one third of the participants indicated they were not familiar with PE student achievement in any of the four domains. Table 5.4 (below) shows the percentage of participants who indicated they did not know in the four domains.

Table 5.4

General Knowledge of Student Performance in LSA (item 10)

	Don't know
Reading	145(28.4)
Math	121(23.7)
Science	168(32.9)
Problem-solving	183(36.1)

The distribution of participants' responses to item 10 was analyzed based on the three grouping variables thought to influence participants' responses. For example, it is likely that parents with children in school would be more knowledgeable about LSAs. Also, participants with higher education would be more knowledgeable about general trends and issues in education. Lastly, it is possible that participants' cultural affiliation, particularly Newcomers to PE, might influence participants' knowledge about LSAs. A cross tabulation of the three independent variables and responses to this item are presented in Table 5.5, below.

Table 5.5

Percentage of Participants who 'don't know' x Independent Variables

	Reading	Mathematics	Science	Problem-Solving
Parental Status				
Parents	27.8	23.1	33.3	36.1
Non-parents	28.7	24.4	32.4	36.1
Educational Attainment				
<=Grade 12	32.0	28.0	38.0	34.0
College/Apprenticeship	43.6	37.3	49.1	53.6
Undergrad.	23.4	18.6	24.8	28.7
Graduate (Master)	22.6	18.2	29.3	34.1
Graduate (PhD)	17.9	17.9	21.4	21.4
Cultural Affiliation				
French	35.1	27.8	37.8	43.2
Aboriginal	23.5	23.5	23.5	29.4
Newcomer	37.3	32.2	37.3	39.8
Local	24.4	19.6	30.6	34.4

Based on the data presented in Table 5.5, the number of parents and non-parents who acknowledged not knowing anything about PE students' LSA scores was about the same for all four domains. However, the are fewer number of both parents and non-parents in the domain of mathematics (i.e., more participants believed they knew about students' mathematics achievement).

Item 15 contained two sub-items (i.e., q15a and q15b) that also explored participants' knowledge about students' performance on LSAs. This item asked participants to compare PE student achievement in relation to other provinces; specifically, the national (15a) and international (15b) assessments using a three-point scale of "below the provincial average", "about the same" and "above the provincial average". The majority of participants, 378 (78.6%), indicated that PE students scored below the average on the PCAP. Similarly, the majority of participants (289, 81.9%) believed that PE students were scoring below the

average on the PISA. Table 5.6 presents detailed information about participants' response to Item 15.

Table 5.6

PE Student Performance Compare with Students in Other Provinces on National (PCAP) and International (PISA) Tests (item 15)

	Below the provincial average	About the same	Above the provincial average
National Assessment (PCAP)	378 (78.6)	97 (20.2)	6 (1.2)
International Assessment (PISA)	289 (81.9)	79 (16.6)	7 (1.5)

Participants' responses were examined based on the three independent variables by using a cross tabulation (see cross tabulation Tables C.4 to C.6 in Appendix C) to examine relationships based on parental status, educational background, and cultural affiliation. To avoid empty or cells less than 5 responses in the cross tabulation table, participants' responses were grouped into correct response (about the same) and incorrect response (above or below the provincial average).

With respect to participants' parental status, approximately the same percentage of parents and non-parent were correct on their PCAP knowledge (Parents: 18.9%; Non-parents: 21.7%). Both parents and non-parents were more knowledgeable about PE's student achievement on the PISA (Correct Responses - Parents: 82.9%; Non-parents: 81.1%). Table 5.7 presents the detailed data of the distribution of participants' knowledge about PCAP and PISA according to their parental status.

Table 5.7

Participants' Distribution (by percentage) of Their Knowledge of PE Student Performance Compared With Students in Other Provinces on the National and International Tests According to Parental Status

	Parents	Non-parents
National Assessment (PCAP)		
Incorrect Response	197 (81.1)	180 (78.3)
Correct Response	46 (18.9)	50 (21.7)
International Assessment (PISA)		
Incorrect Response	42(17.1)	42 (18.9)
Correct Response	203 (82.9)	180 (81.1)

Next, the relationship between participants' educational background and PCAP knowledge was examined. Similar to the above, there was considerably more participants in each educational level who had an incorrect understanding of PE's students achievement on the PCAP and this relationship was not statistically significant ($X^2(4,442)=9.358$, $p=0.053$, phi=0.146) and this pattern also applied to the PISA distribution of responses. The statistical significance for this cross tabulation could not be reported because there was less than 5 responses in one of the cells.

Table 5.8

Participants' Distribution (by percentage) of Their Knowledge of PE Student Performance Compared With Students in Other Provinces on the National and International Tests According to Educational Attainment

	National Assessment (PCAP)		International Assessment (PISA)	
Educational Attainment	Incorrect Response	Correct Response	Incorrect Response	Correct Response
<=Grade 12	37 (74.0)	13 (26.0)	13 (26.5)	36 (73.5)
College/Apprenticeship	73 (70.9)	30 (29.1)	25 (24.3)	78 (75.7)
Undergraduate	116 (84.1)	22 (15.9)	18 (13.3)	117 (86.7)
Graduate	126 (84.6)	23 (15.4)	23 (15.2)	128 (84.8)

Lastly, participants' understanding of PE student performance on the PCAP and PISA was examined by their cultural affiliations. Following the same pattern, approximately 20% of participants had an accurate understanding of PE students' performance on the PCAP (see Table 5.10). With respect to participants' understanding of students' PISA achievement, similar to the other PISA analysis, the majority of participants were knowledge about students' PISA achievement with scores ranging from 77.0% to 83.3% accurately identifying that PE students scored below the national average (see Table 5.10). Given the presence of a few cells with responses 5 or less, it was not possible to report statistical significance.

Table 5.10

Participants' Distribution (by percentage) of Their Knowledge of PE Student Performance Compared With Students in Other Provinces on the National and International Tests According to Cultural Affiliations

	National Assessment (PCAP)		International Assessment (PISA)	
Cultural affiliation	Incorrect Response	Correct Response	Incorrect Response	Correct Response
French	26 (76.5)	8 (23.5)	7 (20.6)	27 (79.4)
Aboriginal	12 (75.0)	4 (25.0)	3 (18.8)	13 (81.3)
Newcomers	84 (78.5)	23 (21.5)	24 (22.4)	83 (77.6)
Local	227 (80.8)	54 (19.2)	46 (16.7)	229 (83.3)
Other	12 (80.0)	3 (20.0)	1 (6.7)	14 (93.3)

Perceptions about LSAs

The next items surveyed participants' perceptions about LSAs. The first three items (Item 11a – c) specifically focused on perceptions about the importance of provincial, national, and international LSAs in providing information about PE student performance. The next group of 12 items (13a – f and 14a – f) represent the scale designed to measure participants' overall perceptions towards LSAs. These perception items combined with the knowledge items in the above section provide insight about the effectiveness of the

accountability role of LSAs in that being knowledgeable about students' LSA scores is the first indicator of the effectiveness of this role. A second indicator is when participants believe that LSAs has some significance to them.

For each of the three levels of LSAs, the majority of participants believed that LSAs at each of the three levels of implementation were important (provincial: M=1.84, SD=0.956; national: M=1.79, SD=0.912; international: M =1.94, SD= 0.966). To examine whether perceptions differed based on parental status, educational attainment, and cultural affiliation, cross-tabulations were used. To minimize the presence of cells with 5 responses or less, participants with less than grade 12 were grouped with the grade 12 participants and at the other end of the scale, participants with Master's and PhDs were grouped together. There was a statistical significant difference between parents and non-parents where almost double the number of parents indicated that provincial (X^2 (3, n=428) = 15.61, p = 0.001, Craver's V = 0.191) and national (X^2(3, n=433) = 13.75, p=0.003, Cramer's V=0.18) LSAs were important having a small to moderate effect size. There was no statistically significant difference based on parental status and perception about the importance of the international assessment where there was more non-parents who believe the PISA was important (X^2(3, n=412) = 6.21, p=0.102, Cramer's V=0.123). When comparing participants' level of education and importance of the three LSAs, the majority of participants' (at least 70%) at each level indicated that LSAs were important (responded with a 1 or 2 where 1 was anchored with very important). Given the few responses in the low end of the scale (3 or 4), there were a number of cells less than 5 hence the low number violated the assumption underpinning the Chi-square test. The bar char shown below (Figure 5.1) displays how perceptions of the importance of each LSA declines with educational attainment.

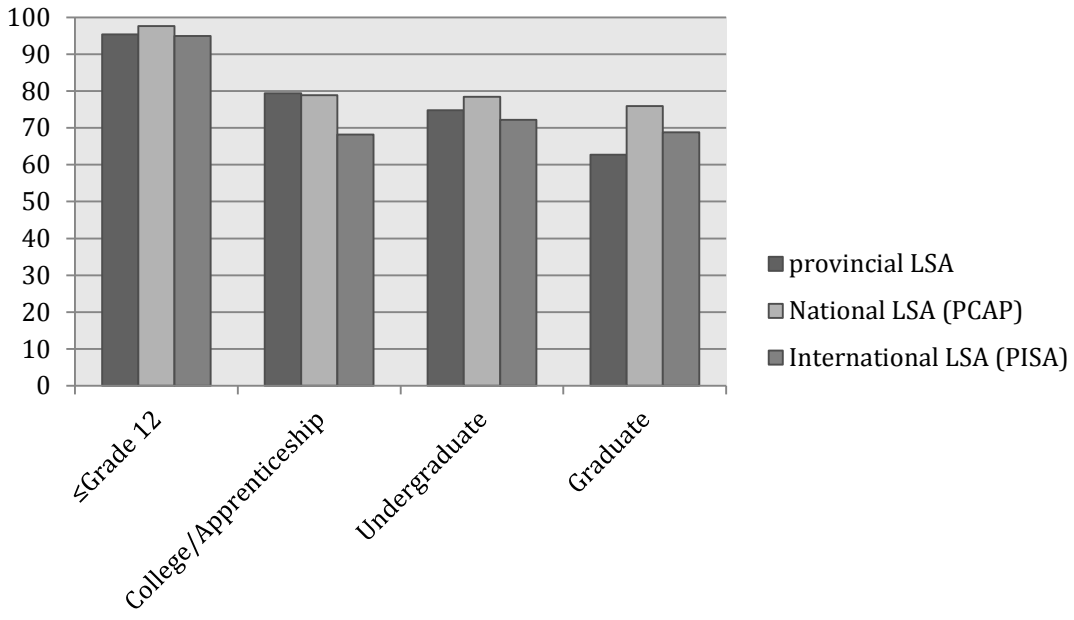


Figure 5.1. Educational Attainment x Perceptions of the Importance of LSAs at the Provincial, National, and International Level

When participants were grouped by cultural affiliation, the majority of participants (60% or higher) indicated that LSAs at each level were important (responded with a 1 or 2 where 1 was anchored with very important). Figure 5.2 below, displays the differences in perceptions of the importance of each LSA. The figure shows that the percentage of Newcomers, who reported that LSAs were important, was slightly higher than the percentage of participants in the other three groups. Since minimum cell count was violated, the significance of this relationship could not be reported.

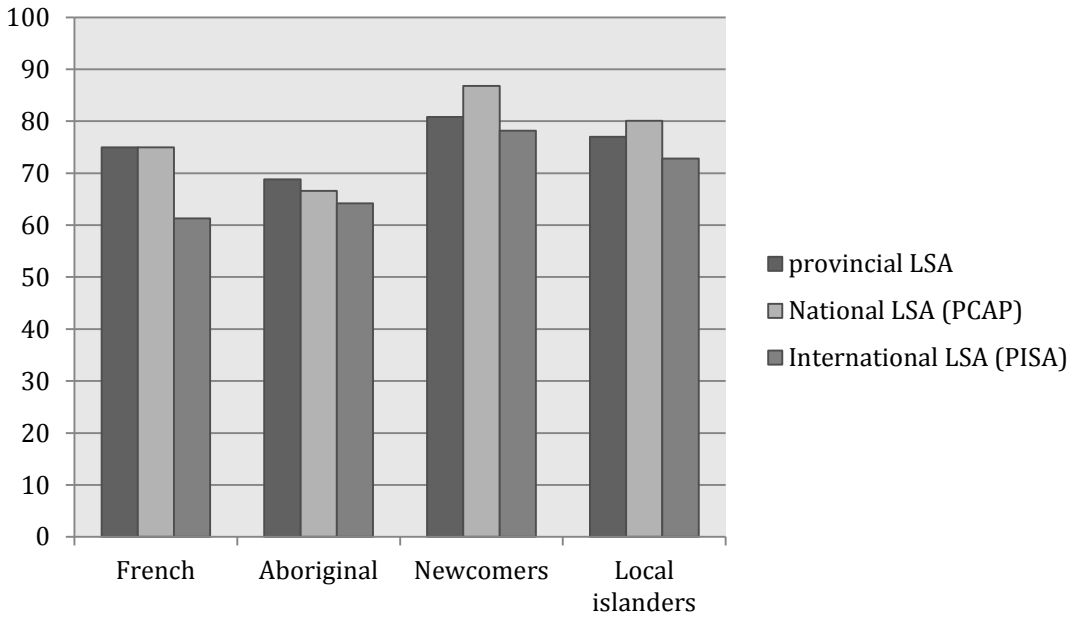


Figure 5.2. Cultural Affiliation x Perceptions of the Importance of LSAs at the Provincial, National, and International Level

To explore the characteristics of non-believers (i.e., participants who responded with a three or four on the four-point scale where 4 represented not important), this group were selected and their demographic characteristics were analyzed.

There were 99 participants who were negative about provincial assessment (who rated "3" or "4" in the scale). 67 (68.4%) participants were born after 1964, 77 (79.4%) were female, 42 (42.4%) participants in this group were parents, 73 (76.1%) were employed (i.e., 54 were employed full time; 9 were employed part time, and 10 were self-employed). Among those employed participants, 54 out of 73 (58.7%) were involved in education, law, and social services related occupations, 10 (10.9%) were working in business related fields. As for the education background, the majority of participants held a bachelor degree (32 out of 99) or 33 out of 99 held a Master's degree. Regarding the culture affiliation of these participants, the majority (58 out of 99) were from the local community.

Perceptions on Moving to a Better Performing School

To further explore the extent to which the accountability purpose of LSAs was effective in PE, participants were given an extreme context in which they were asked to indicate whether they would move their child to a school that had better provincial test scores, if such a move was indeed possible (Item 12). Although there are a number of factors that would influence parents' decision to move schools, this item is another indicator of the extent the accountability framework is working. If parents were seriously concerned about their children's learning as measured on the PISA, they might choose to advocate for an improvement in teaching or learning resources (in hindsight, this option should also have been posed to participants). Alternatively, parents may choose to move their child(ren) to a better performing school. This item was posed to parents using a dichotomous scale (move to a new school or not move). Responses were evenly distributed between those who would move to another school and those who would not (Move to new school: 243, 49.5%; Not move: 248, 50.5%)

In taking a closer look at the characteristics of participants who indicated they would not move schools there was no characteristics that stood-out from the sample of participants., The distribution of participants who would not move were female (188, 77.7%) (but the overall majority were also female). As for the age range, 4 (1.6%) were above 70 years of age, 79 (32.0%) were from 51 to 70 years old, and 164 (66.4%) of the participants were under the age of 51 which in this age category, are more likely to have a child still in school. Regarding the employment status, there were 194 (80.8%) participants who had jobs, with 146 (60.8%) full-time employed, 22 (9.2%) part-time employed, and 26 (10.8%) self-employed. Among those participants who were employed, nearly half of them were

working in education and social work related fields. As for participants' educational background, less than a quarter (51, 22.0%) of the participants had college/apprenticeship certificate, 75 (32.3%) held undergraduate degree, 76 (32.8%) hold master degree, 9 (3.9%) finished their PhD program. As for the cultural affiliation, over half of the participants (149, 65.4%) were local residents on PE, and nearly one third of the participants were Newcomers (44, 19.3%). The characteristics of participants who indicated they would not move their child(ren) could primarily be described as females, people who were employed, and local islanders.

Perceptions towards LSA

The descriptive analysis of items (frequency, percent, mean, and standard deviation) representing the scale measuring perceptions towards LSAs is presented in the section below. Following is the inferential statistical analysis which reports on the factor analysis and the analysis of variance used to examine differences between the scale and the three independent variables (i.e., parental status, educational attainment, and cultural affiliation).

Descriptive summary. A 6-point rating scale was used where six represented the negative end of the scale and one represented the high end of the scale or those who held positive or strong perceptions towards LSAs for the first set of six items. The scale for the second set of six items was similar but the negative end of the scale was anchored with *strongly disagree* (position 6 on the scale) and the positive end of the scale was anchored with *strongly agree* (position 1 on the scale). This scale contained an option for participants to indicate they did not know anything about what was being asked.

Drawing on the mean and standard deviations, participants were not overly convinced that LSAs told them about whether the education system was working (Item 13a: M=2.57,

SD=1.58) and that children have the necessary skills for success (Item 13b: *M*=2.91, *SD*=1.60). Similarly, participants were not certain that LSAs told them about children's success (Item 14a: *M*=3.1, *SD*=1.58) or about the quality of schools (Item 14c: *M*=3.04, *SD*=1.65). Participants were more positive that the LSAs were used to compare PE with other countries (Item 14d: *M*=2.08, *SD*=1.20). On both items that related to the quality of teaching, participants were slightly more negative (Item 13f: *M*=3.31, *SD*=1.67 and Item 14e: *M*=3.31, *SD*=1.7). When surveyed about the relationship between businesses and LSAs, the majority of participants did not believe that LSAs provided information about the success of businesses (Item 13e: *M*=4.33, *SD*=1.66). When probed further about this relationship, about a third, 175 (34.9%), did not know whether business looked at PE's student achievement on LSAs (Item 14f). For participants who provided a response to this item, the majority did not agree (i.e., selected 5 or 6), that businesses looked at the LSA scores for PE (*M*=3.29, *SD*=1.68). Also noteworthy, was approximately a third of the participants who indicated they did not know whether people considering moving to Canada would look at LSA scores (Item 14b). Moreover, 34.9% (175) of the participants indicated that they did not have any awareness that business thinking about moving to Canada would look at PISA scores (Item 14f). The characteristics of these 199 people are as follows: the majority of the participants were female (158, 80.6%), 137 (69.2%) of them were at or below the age of 51, more than half were from the local community (113, 61.7%), and nearly one quarter of them were Newcomers (49, 26.8%). See Appendix C for a complete copy of the descriptive statistics.

In sum, LSA data are indicators of the wellbeing of the education system and economic prosperity (Hayward, Pannozzo, & Colman, 2007). However, many participants in this

study do not see the link between LSAs and businesses (economic prosperity). The relationship between the education system, as explored through students' success and the quality of teaching, is stronger but not overly strong. For example, the majority of participants believe that LSAs tells us a lot about whether the education system is working but participants are not overly convinced that LSAs tells us a lot about student skills or success, or the quality of teaching.

Factor analysis. The 12 items measuring perceptions towards LSAs were subjected to a principal components analysis using SPSS version 23. Prior to performing the principal component analysis, the suitability of data for factor analysis was examined. Inspection of the correlation matrix revealed that all coefficients were greater than 0.3. The Kaiser-Meyer-Olkin value was 0.924, exceeding the recommended value of 0.6 (Kaiser 1974) and Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance, indicating the factorability of the correlation matrix.

The principal component analysis revealed the presence of two components with eigenvalues exceeding 1, explaining 60.6% and 20.7% of the variance. An inspection of the scree plot revealed a clear break after the first component hence one component was retained for further analysis. All items loaded substantially on the one component. This analysis confirmed the one-dimensional nature of the scale measuring perceptions towards LSAs.

Table 5.12

Factor analysis, component matrix

Item	Component
The tests tell us about the quality of our schools	0.858
The tests tell us about the quality of our teachers	0.858
... the education system is working	0.852
... teachers are doing their jobs	0.848
The tests tell us about children's success in school	0.836
... students have the skills to take part in the world today	0.823
... students have the skills to be successful at college or university	0.807
... children can solve problems in their own lives	0.788
The tests are looked at by people thinking about moving to Canada	0.710
... businesses are successful on the island	0.674
Businesses thinking about moving to Canada will look at the test scores	0.649
The tests compare us to other countries	0.572

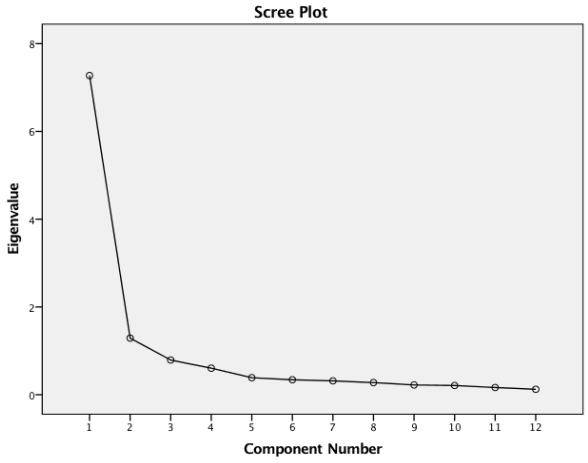


Figure 5.3. Scree plot

One-way ANOVA. A one-way ANOVA was used to determine whether there were any significant differences in the scale representing participants' perceptions towards LSA, as measured using the compressed scale (all items were summed together) presented above, and the three independent variables: Parental status, educational background, and cultural

affiliation. Recalling, educational background contained four groups: grade 12 diploma or less, college/apprenticeship certificate, undergraduate degree, and graduate degree (Master's or PhD). The other independent variables did not have any outlier populations. The ANOVA revealed that there were no significant differences at $p<0.05$ level in any of the three independent variables (see Table 5.13 below)

Table 5.13
One-way Analysis of Variance of Total Perspectives on LSAs in Parental Status, Educational Attainment, and Cultural Affiliation

	df	*SS*	*MS*	*F*	*p*
Parental status					
Between the groups	1	1.667	1.667	0.867	.352
Within groups	390	750.202	1.924		
Total	391	751.869			
Educational attainment					
Between the groups	4	14.681	3.670	1.991	.095
Within groups	363	669.097	1.843		
Total	367	683.777			
Cultural affiliation					
Between the groups	4	4.455	1.114	0.570	.685
Within groups	366	715.089	1.954		
Total	370	719.544			

Note. Participants who have less than Grade 12 education attainment were excluded due to the small sample.

Although there was no significant difference when examining the three potential influencing factors as parental status, and cultural affiliations, some trends were evident. The following sections reports on participants' perceptions about LSA regarding three aspects: Firstly, LSA as an indicator of education system accountability; secondly, LSA as evidence of education gatekeeping; and lastly, LSA as a predictor of children's global competency and local economic prosperity.

Perceptions related to educational accountability. Participants were surveyed about LSA as an indicator of the accountability of the education system in Items 13a, 13b,

13f, 14a, 14c, and 14e. Participants who rated 1 (tells us a lot or strongly agree) and 2 in the scale were considered to hold a positive perceptions towards LSAs. Over half of participants (275, 54.9%) believed that these LSA results informed the public about the effectiveness of the education system (*M*=2.57, *SD*=1.583) (Item 13a). However, for the international assessment PISA (Item 14c), only 190 (37.8%) participants agreed that this test would provide information about the quality of schools (*M*=3.04, *SD*=1.651). These responses showed that for some people (not the majority) LSAs indicated the effectiveness of the education system in the eyes of the public, especially for provincial and national assessments. When focusing on the use of LSAs in a general context in telling us about the quality teachers, 195 participants (39%) believed LSAs are indeed indicators of teaching quality (13f: *M*=3.13, *SD*=1.67). However, when the similar question was posed related to the information the PISA provided about the quality of teachers, only 150 participants (30%) believed the PISA was an indicator of the quality of teachers (14e: *M*=3.31, *SD*=1.706). This finding reveals that the public has little confidence in LSA's predictive utility in reporting on the effectiveness of teachers, especially for the PISA. Further inquiry was explored to examine the public's perspective about students' success in school. 45% of participants believed that, in a broad sense, LSAs were indicators of students' ability to be successful in post-secondary (13b: *M*=2.91, *SD*=1.60) while only 33.3% of participants would agree that the PISA had a similar function when this statement comes to certain international assessment (Item 13a: *M*=3.10, *SD*=1.586).

Perceptions related to children's academic success and global competency.

Participants' perceptions towards LSA as an indicator of children's academic success can be inferred from Items q13c, q13d, and q14d. These three items examined participants' perceptions related to utility of LSA in predicting students' future life skills and global competency.

In terms of using LSA achievement as a predictor of children's ability to solve real-life problems, 146 participants (29.4%) believed (i.e., selected 1 or 2) that LSAs provide information about children's problem solving abilities (Item 13c: *M*=3.42, *SD*=1.65). A similar number of participants (142, 28.5%) reported (i.e., selected 5 or 6) that LSAs provided little information about children's problem-solving skills for the same item. A large number of participants (210, 42.2%) were quite neutral about this statement (*M*=3.42, *SD*=1.65).

In terms of children's global competency (Item 13d), one third of the participants (159, 32%) strongly agreed or agreed that LSA data provides information about whether students' have the skills to participate in today's society. Most participants were less positive responding at the opposite end of the scale indicating they did not perceive LSA data as useful in providing such information (Item 13d: *M*=3.37, *SD*=1.64). With respect to the PISA, over half of the participants (300, 70.8%) responded with a one or two that the PISA data is a good indicator to compare PE with other countries (Item 14d: *M*=2.08, *SD*=1.299).

Perceptions related to local economic prosperity. Items 13e, 14b, and 14f surveyed participants' perceptions about LSAs as an indicator of local economic success along with attracting people and businesses to PE. Over half (54.6%, rated 5 or 6 where 6 is strongly disagree) of the participants did not believe that there was any association between student

performance and the development of the local economy (Item 13e: *M*=4.33, *SD*=1.66). When this question was asked in the section focusing just on the PISA, approximately one third of the participants indicated they did not know whether PE's PISA scores were viewed by people considering moving to Canada (167, 33.3%) or business (175, 34.9%) moving to PE. Those who knew more about the PISA indicated that people moving to Canada would look at the PISA scores (Item 14b: *M*=2.79, *SD*=1.655), but that it would not indicate business success (Item 14f: *M*=3.29, *SD*=1.676), with 158 (31.5%) and 118 (23.5%) in agreement respectively.

Interview Findings

Four in-depth interviews were conducted to expand on the questionnaire findings and further explore aspects of LSA culture and practices. As noted in the methods section, the interviews were conducted with a Provincial Achievement Specialist, High School Principal, Representative of the Bioscience Business Sector, and Representative from PEI Home and School Federation. This section begins with participants' feedback on three general questions, followed by participants' own perceptions and values about LSAs as it relates to their personal expertise and experiences. This section on the interview findings is concluded with a summary of the key findings.

General Question 1: Share what you know about PE student performance on the most recent provincial, national and international assessment.

Three out of the four participants indicated they knew that PE students were poor performers on the national and international assessments, especially in the domains of mathematics and science. These three participants articulated that PE students were always situated at the bottom compared to their national and international peers. With respect to PE

students' improvement on the 2012 PCAP test, the Provincial Achievement Specialist noted that, "Although there was a slight improvement in PCAP from the time before, not enough …, we still have a lot of students at the lowest level." This perspective was affirmed by the Representative of the Bioscience Business Sector and the High School Principal. In contrast, the Representative from the Home and School Federation expressed an opinion different from these three participants. He noted that PE students were not performing so poorly since their average score had been above the OECD level before 2006 and the OECD's average was quite high. Unfortunately, this participant may not be aware that some of the European countries like Germany and France were improving their rankings to compete with other regions like Macao-China, whereas PE ranking had slipped through four cycles of the assessments.

Additionally, the Home and School Federation representative indicated that educators and researchers should not simply rely on one year's data to draw conclusions of whether students have made progress or not. He held a cautious attitude about using the PISA results to guide any initiatives and plans to change the system. However, upon examining the PISA results from 2000 to 2012, PE students have always been situated at the bottom of the scoring scale. This is obviously not one year's worth of data displaying student achievement hence this interview participant has some ill conceived perceptions of LSA. Although there are new initiatives, such as the new Grade 9 mathematics curriculum launched in 2010, it was initiated almost ten years after PE first joined the PISA in the domain of mathematics. During these ten years, three cycles of the PISA data have been provided as evidence that PE students poorly performed in mathematics both nationally and

internationally.. This accumulative data should guide policy reform and influence changes in the provincial education system.

When focusing on provincial assessment scores, the Provincial Achievement Specialist noted that PE students showed a slight improvement in grade 3 and 6 reading, but the writing scores decreased partly because the test used a different method to assess students' writing skills. Specifically, the provincial assessment has adopted multiple choice questions and open-response questions, which lower grade teachers would not use in classroom. With respect to Grade 9 writing assessment, the test measured a different form of writing, which was very demanding for students.

In terms of the grade 3 and the newly piloted grade 11 mathematics assessments, two of the participants believed that the results were not satisfactory. The High School Principal described students' performance on the grade 11 LSA (pilot) of mathematics as terrible. She explained that the LSA average in the 521B course, which is a pre-calculus course and the most challenging of the grade 11 mathematics courses, was 40% lower than students' course averages. She also added that this misalignment between students' classroom and LSA achievement rattled students who thought they were good at mathematics. Subsequently, these disappointing scores influenced some students to alter their ideas about applying for various post-secondary programs.

General interview question 2: What do you think are the applications or uses of provincial assessment (common assessment), national assessment (PCAP), and international assessment (PISA) in PE?

The common theme from all four participants was that LSAs should inform instruction at the school level and that it provides a picture of the wellbeing of the entire education

system. One of the participants strongly believed that the standardized nature of LSAs was the only way to objectively determine whether student performance was improving or not (the Representative from Bioscience Sector of Business). Three of the participants believed that the provincial assessments were effective instruments used by policy makers to launch new initiatives. Provincial LSAs were also believed to provide teachers and school leaders with a better idea of the starting point of their teaching. The High School Principal stated:

> The data we got from doing the writing [assessment], was impressive. [That assessment] really informed our teachers about where the strengths are, where the weakness are, and what they [teachers] needed to put into place and move them [students] forward. It [LSA data] should help us understand not to teach the things that the kids already have in place and focus on the ones that would help them move forward.

The Representative from Home and School Federation echoed this with:

> The Common Assessments are definitely used by the Department to make changes, in their mind, positive changes to the education system. As a result of the PE Common Assessments, the Department as well as the School Board has introduced new initiatives including literacy coaches. So the School Boards have done a lot to ensure that the students are at the right level by reading at grade 3, which is directly drawn from the common assessments.

When discussing the national and international assessments, two of the participants (the Provincial Achievement Specialist and the Representative from Home and School Federation) reported that they were aware of what that the Department of Education of PE

had been doing with the PISA data. Furthermore, the Provincial Achievement Specialist admitted that the school system had just started to recognize the importance of student achievement on the PCAP and PISA. The Department has brought representatives from the CMEC to speak to the principals and to test the teachers to see whether they can in turn teach the kinds of skills and strategies that students need to perform well in the next round of PISA and PCAP. Teacher can also use these practice tests to help their students to prepare the actual tests. However, no data from teachers and students was required to be collected for further analysis. However, the Achievement Specialist acknowledged that PE students have scored poorly on the PISA since 2000 and wondered, like myself, why it has taken more than a decade for the related educational organizations to realize the importance of these tests and offer resources at the school level. Interestingly, this participant has worked at the Department of Education since 2006 as a Provincial Achievement Specialist yet she has deferred the responsibility of addressing gaps in student achievement to related educational organizations.

Another theme drawn from the interview data was the absence of discussions from all participants about student achievement on the PCAP. Although PE students have made progress on the PCAP, especially in the domains of mathematics and science, the Achievement Specialist did not mention any initiatives stemming from this data set. Likewise, the representative from the Home and School Federation reported that he was not aware of any initiatives using PCAP data both from the Department of Education or from his daughter's school. It is possible that these and the other two interview participants had merely focused their responses on how the PISA data was used or it is possible that the

international aspect of the PISA overshadows the PCAP. In hindsight, the participants should have been prompted to speak more specifically about the uses of the PCAP data.

The Representative from the Home and School Federation also spoke from the perspective of a parent with a child in grade 9 mathematics. He described how his daughter's mathematics teacher determined his daughter's final standing in the course by incorporating part of her achievement on the provincial grade 9 assessment of mathematics with her teacher derived score. He indicated that after receiving the report card from school, he made some inquiries with other parents in the neighbourhood and found that despite the government's directive for teachers to incorporate 10% of students' grade 9 provincial assessment of mathematics score with students' final grade, not all the schools and teachers followed this directive, a finding which was reflected in Miller's study (2013).

The Representative from Bioscience Sector of Business strongly expressed his disappointment with the LSA implementation in PE public school systems. He commented that PE students have been struggling at the bottom of national and international assessment through the past decade. However, there has been little change. He quoted Einstein's saying "if you keeping on doing the same thing and expecting different results, you must be a fool" to urge the provincial government to make dramatic improvements in student learning outcomes.

General interview question 3: One of the purposes of standardized tests is to measure the functionality of the education system. Discuss whether this is a good indicator of the education system.

Although there are many differences between formats, styles, and purposes of the three LSAs, all participants agreed that LSAs were key indicators for educational accountability.

The Achievement Specialist and the Representative from Bioscience Business Sector believed fully in the importance of LSAs at all three levels, while the High School Principal favoured using provincial LSAs for accountability purposes. She commented, "They are created by the teachers in our system. So, I have faith in their ability to measure what they are supposed to measure."

With respect to the national and international assessments, the Home and School Federation Representative expressed concerns about relying too heavily on making decisions based on student achievement from one LSA. In comparison, the Achievement Specialist was also a little hesitant to acknowledge that these assessments had actually been used as indicators of accountability, but thought that the current focus on the importance of these assessments would translate into better indicators of accountability. Although all interviewees believed LSAs were good indicators of educational accountability, the extent LSAs are actually holding those responsible for educating the public accountable was less optimistic.

Perceptions related to accountability of the education system. The Achievement Specialist indicated that at the current time, there were no graduation standards or outcomes that articulate graduating student competencies. PE's low dropout rate is a direct result of this absence of standards, since all students in Grade 12 receive their graduation diploma (Achievement Specialist). According to the Achievement Specialist, this absence of standards was the driving force behind the Department of Education's initiative to introduce a literacy assessment in Grade 10. This LSA would subsequently ensure that students who pass the assessment would have attained a basic level of literacy.

The Representative from the Bioscience Business Sector also expressed similar concerns about the education system's wellbeing. He agreed that the difficulty in the system was that a lot of students who go through the education system in PE would never really reach the competency level needed to get into a good university in Canada or in other countries. He also elaborated that if the education system has poor undergraduates, it is because the primary education is poor. Students without a good foundation are not going to do well at the university level.

> There is a policy to push them through. Because they don't want to mentally challenge them, so, the students don't meet the requirements. To me, that is very short-sighted, and it is not benefiting the students. Because if the students get to the end of the high school, they can not read, they can not write, they can not do math, even though they graduate, then how successful would they be?

The high school principal and the representative from Home and School Federation held different perceptions about the new provincial assessments in Grade 10 and Grade 11. The high school principal elaborated on the technical issues of the literacy test and the poor student performance at her school. She was more concerned about the quality of the teaching staff to help students make it through the new assessments. She further explained that if the tests are based on the provincial outcomes, and the students get high marks, then it should mean that the students are well prepared for their next step from Grade 12 onward into colleges and universities. She was sceptical, however, that LSAs can be a single indicator of student potential to succeed. She noted the importance of which the student's attitude, mental health, work ethic, and motivation for learning played in determining student success.

The Representative from Home and School Federation expressed that he was not so sure about the new assessments. His main concern was that the tests were too late to measure student's opportunity to attain post-secondary education. He indicated that since the assessments are measuring the system, not individual students, the interventions needed to be established from the earliest stages of primary education rather than only at the end of the high school. He considers that the new assessments are regarded as a kind of exit tests for high school graduation, rather than for simple diagnostic purposes. For the Grade 10 literacy assessment, if the students failed the first time, they are still eligible to take the assessment in Grade 11 and Grade 12 before they graduate. That is two more years for student to improve and meet the standard.

Perceptions related to children's academic success and global competency. Based on the poor performance that PE students have in the three levels of LSAs, participants were asked to postulate the main cause of this situation. Participants were concerned with curriculum, the quality of teaching staff, the policy and the culture that parents hold for education. The Representative from Bioscience Business Sector expressed his concern of the quality of teaching staff and their capability to carry the curriculum outcomes. Two of the participants made comments on the policy of the snow days. Last winter (2014-2015 school year), there were 13 snow days where the school was closed for safety reasons. However, the schools did not make-up the lost instruction time. That period was almost three weeks of school time, moreover, the last week of every semester is always not teaching week but play week (the Representative from Bioscience Business Sector). The lack of instructional time at school is one of the main concerns the participants held for student academic success.

The last reason is about parents' attitude and the public culture. Two of the participants indicated that there is not an education culture in PE. The representatives from Home and School Federation commented that: " We do not have a culture thing here for many youth. In PEI, education is not considered as a priority. For many parents in PEI don't find education a priority." The representative from Bioscience Business Sector echoed with: "Culture on this Island, education is not a priority. That is because a lot of history and culture are primarily fishing, farming, they don't believe either in education." Both representatives from Home and School Federation and Bioscience Business Sector explained that sporting events and extracurricular activities take priority over instruction, which leads to high absentee rate in schools. The representative from Home and School Federation also noted that a second reason for that high absent rate is that students do not have a clear sense of what they are going to do in their career path and they do not have a focus.

Perceptions related to local economic prosperity. When referring to the knowledge and the skills that students need to be equipped for the future work, strong basic knowledge in literacy, mathematics, and science are considered to be the very fundamental capacity of future workforce by participants (OECD, 2003). Other than this basic knowledge, as noted by two participants (representatives from Home and School Federation and Bioscience Business Sector), problem solving and communication skills are also important skills that students should have in order to take an active part in their work and live, thus, to enable themselves more "knowledge" and "able". These two vital skills were considered important factors that decide whether student could get a job or how far the student can keep on his/her career. The poor results from LSAs discouraged employers' confidence of opening new

business in PE because they could not find qualified talents for the job. The high unemployment rate of young people between 18-30 is a warning sign for economic development (representatives from Bioscience Business Sector).

Special questions for the provincial assessment specialist. Eight questions were posed to the Achievement Specialist from the Department. Question 1 was about the alignment of the classroom test to the provincial assessments in the domains of literacy and mathematics. With respect to the subject of mathematics, she indicated that although classroom tests were based on the curriculum, she was not sure that the classroom tests were as well as developed as provincial assessment. In terms of the difficulty levels of these two types of assessments, she commented that before the Department had launched the projects and workshops for mathematics teachers, the knowledge of how to create correct levels of classroom assessment was not there for mathematics teachers. Although the training has been brought to more and more classrooms and grades, she was not sure that teachers were bringing the knowledge and skills back into their classroom to influence their instructional and assessment practices to be more aligned with provincial assessments. As far as the literacy assessment, she stated teachers were not really questioning the students and were not mindful of the types of questions they needed to be posing to their students. She further elaborated that after examining students' poor performance on the PCAP and the PISA in the reading comprehension part of the literacy test, especially in the nonfiction subdomain, the Department team looked at classroom instruction in nonfiction and found that even though teachers had the resources in the classroom, they never taught nonfiction. According to her understanding, the teaching approaches of fiction and nonfiction are rather different, and many teachers refuse to teach nonfiction. The lack of any instruction and assessment on

nonfiction type of reading comprehension could probably form part of the cause of poor student performance on the PCAP as well as on the PISA.

The next question about provincial assessment standards was raised. The Achievement Specialists described the literacy assessment, reading and writing, as separate assessments. In terms of reading, the provincial assessment has a cut score, which is basically the percentage of students that are determined to pass. The cut scores are developed using statistical process, and are set at each grade level. It is interesting to find that the cut scores are not the same number of percentage of students reported on the government reports released publically. Take the Grade 3 reading test for example: if the cut score was 68, the provincial government would report that there were 71% of students at or approaching the standard, actually, there was always 3% of students who were below the cut score. Those 3% of students were not meeting the standard. The government added the approaching number on the cutting score just to make it look better (Achievement Specialist).

As for the writing part, the Achievement Specialist believed that "student performance was shown in four categories of experiencing difficulty, approaching expectations, meeting expectations and strong performance" (Achievement Specialist, 2015). In each of the categories, three writing subdomains were presented for markers to check. In terms of mathematics, the participant did not know either the cut score for mathematics, nor the description about each standard of the math assessment for each key grade.

The third question was about the inclusive plans on provincial assessment for PE students. With the trend towards inclusive education on provincial assessments in other provinces like Ontario and Alberta, there have been initiatives for governments as well as

educators to provide special assistance on LSAs for student with learning disabilities. According to the participant, there was no evidence that PE has similar strategies for students with special needs in completing LSA. There were certain adaptations (e.g., extra time, using scripts) students could use, but the Achievement Specialist noted that some school administrators were not sure about whether or when they could use these adaptions, so the Department people needed to offer assistance to the students themselves. When the new assessment for Grade 10 literacy is implemented, all students would be allowed to use the adaptations since all the adaptations will be automatically programmed in the online testing system. However, scores from students who receive adaptations were not included in the summary of provincial assessment results. Regarding to this practice, the Achievement Specialist explained, as an example, that students could have extra time, however, incorporating their scores with students who do not receive any adaptations would cause invalidity of the assessment. The Achievement Specialist further explained: "If the scene is if I am measuring your reading, and you are over here someone is reading for you and I am not measuring your reading, I am measuring your listening comprehension, right? So, it's two different things. So, compromise its invalidity". The same issue was raised by the High School Principal as well.

Question four examined the Achievement Specialists views regarding the benefit or limitations of using online technology to administer the Grade 10 literacy test. The participant agreed that online technology was a benefit for students, mainly because student life was connected to technologies such as cell phones, tablets, and laptops. Moreover, teachers now use internet technology in classroom teaching. So, students are quite comfortable with that. She expressed some interest in looking for achievement gaps in

reading and writing between genders when using this new online assessment of literacy to see if there were any evidence to show that boys may perform better on online tests than paper-based test, which might possibly reduce the performance gap.

The participant was asked about the financial model of provincial assessment. The participant indicated that the budget was about $2 million (Canadian dollars), which she considered a very small amount. The participant also expressed different opinions from the views presented during the provincial leadership debate about education in April 2015. Instead of the Teachers' Federation's position of cutting provincial assessment funding and put it towards hiring more teachers, the participant strongly opposed that the government could not afford to lose the provincial assessment by simply hiring extra teachers or purchasing more resources. She stated that the provincial assessment created a lot of school improvement, teacher professional development projects and curriculum updates or revisions by improving student scores, teachers' instructional practices and overall school performance with further comments: "The data itself, has probably created almost like resources for teachers. Help you [teachers] be more efficient, more aware, a better instructional leader, which is more important than just to hire another teacher and then still doing things that are not supposed to be doing."

The Achievement Specialist strongly believed and commented: "[the] provincial assessments maintain accountability within the education system; it puts the system in check, which is more important than hiring a few new teachers and equipping resources." In the case of hiring new teachers, the Achievement Specialist explained that the government cut the number of teaching positions from 40 to 12 at the beginning of the school year, which made the situation more stressful for teachers. She also indicated that her team was cut and

now there were only five people overseeing provincial assessment, which frustrated her very much.

Perceptions about the new provincial assessment for Grade 10 literacy and Grade 11 mathematics were inquired about to see to what extent that participant consider these two new assessments would be good indicators of students' post-secondary attainment. The participant indicated that PE had a low dropout rate because before the new assessments were introduced, there was no clear standard; everybody received a diploma from high school. Unfortunately, students found that they were less competent in post-secondary education and their careers as well as were less confident, yet had high marks on their high school report card. The Achievement Specialist commented that this has been an issue and needed to be addressed. The main reason for the introduction of the new assessments was to create standards in literacy and mathematics to guarantee that students are getting the required knowledge and skills to pursue their post-secondary education and future careers. The participant believed that the new assessments were good indicators, since they provided feedback on the skills and strategies that students needed. She also commented that because these tests set clear standards, there would be a lot of scenarios happening, such as stress, fear, and anxiety that both students and teachers would suffer.

Given the poor performance that PE students have consistently demonstrated, the participant was asked whether there were any plans being implemented or to be launched to ensure students received the basic skills a) to solve real world problems; b) to build learning strategies; and 3) to take active parts in the 21st century. The participant restated that there was lots of work to be done in this aspect. The only thing that she was aware of was that the Department was now talking to the principals about the curriculum, the assessments, and

teaching instruction. This participant was concerned about is that change takes a very long time in PE. Working with principals was just a baby step. Other than the time issue, she did not believe that educators would buy into the idea that LSAs are assets to the education system, due to the low value educators and the public held for LSA as a cultural issue, which need to be changed.

The last question posed to the participant was about the implication of provincial assessment, mainly referring to teacher professional development in the aspect of knowledge and skills related to the assessments. The participant was very confident about the growth of teachers' knowledge and awareness in writing instruction and assessment. She said, "We are understanding writing more and more" (Achievement Specialist). Before the first implementation of the grade 3 Language Arts provincial assessment, teachers never used rubrics and did not know about the different forms of writing students needed to learn and practice. This statement echoed with the CMEC's finding of students from PE who were less familiar with rubrics in writing compared their counterparts from other provinces on the PCAP test. As for mathematics, teachers began to change their instructional practices in class after they took part in the mathematics project which was initiated in 2008 by Western School Board, later extended to both of the English and French School Boards. Teachers collected students' data, diagnosed and developed strategies for better instructions. A curriculum, instruction, and assessment coach worked with mathematics teachers to provide the support. Unfortunately, this support did not last a long period (Department of Education, Early Learning, and Culture of Prince Edward Island, p9). However, she acknowledged that mathematics teachers still had a great deal more to learn.

Special questions for the high school principal. There were seven questions about how LSAs influence classroom instruction and school improvement. The first question was about the alignment of the standards of classroom tests with the provincial standards. The participant indicated that teachers in her school check the provincial assessment results and reshaped their teaching plan to cover certain areas in curriculum in order to help students score well in their tests. This was especially the case for the new assessment for Grade 11 mathematics assessment, which counts for 25% of the final grade and appears on the students' transcripts, which creates a strong sense of a high stakes assessment. As for the instruction part, she expressed her concerns about classroom instruction and the chances of professional development for teachers in terms of enhancing their competency to combine key ideas and contents from the provincial assessment into classroom practice. The first issue she saw from the classroom instruction was that the wording on test items was different from traditional teaching practice.

> The way student are taught math in school now is run Chapter 2, which is about quadratic function, we learn quadratic function, we get test on quadratic function. They know math since if they know it is "quadratic function". The comprehensive Grade 11 Math exam what student ability to apply in real life scenarios, that is why students fell down and had bad marks.

The second issue was how teachers' professional development relates (or does not relate) to the new provincial assessment. She elaborated on this issue by describing the case of two mathematics teachers who teach 521A and 521B respectively. When implementing the pilot version of the grade 11 provincial mathematics assessment the principal noted that the 521A mathematics teacher was on the curriculum committee as well as the marking

board, so she knew how to gear her instructions to the weak areas of the curriculum according to the current student ability. This experience turned out to be a huge success since the score differential between her class scores and the scores on the provincial assessment was only 3%. Conversely, the 521B teacher did not receive any opportunity for this type of professional development. The score differential for the 521B teacher was 40%. The principal said that the 521A students were lucky to have their teacher on the committee, since there was no professional development related to the provincial assessment for the average mathematics teacher who delivers the course. "So, you launch the test, but the teachers do not know how to teach it yet, because they had not exposure to it. This happens to Grade 9 math assessment as well." (High School Principal) She commented that the good solid academic students were now questioning whether they should take the advanced course because they scored so poorly on their provincial assessment. Some of them even changed their post-secondary plans based on the result of one provincial assessment that had been run for the first time without any teacher preparation, which she thought was a big mistake.

The second question was about the utility of provincial assessment results for school improvement. With respect to this question, the Principal first explained as the following:

> So in the second semester, what I did, it was obviously that one of the teachers [521A teacher] had more competency, understand the test and how the test was set up and she was also the department head, so, I freed her up to go, well, she had a work period. So she worked a little bit with the other teacher, and tried to gently to get him to, she helped him [521B teacher] so that the kids could pick up what they miss. Some of the math teachers volunteered to give extra help, they came along on Saturdays and studies the provincial assessment, so everybody kind of realized that the kids are still lacking

skills and they need to pick them up quicker. So, the math people kind of wrapped around.

With respect to other provincial assessments in grades 6 and 9, as a principal, she created student packages containing student profiles that her high school teachers could use to plan their instructional practices and present lessons that were not too overwhelming for weak students, nor too boring for the stronger students. However, there is no evidence that high school teachers would actually change their instructional practices based on the data they received from provincial, national, and international LSA, since as noted before, from the PCAP 2010 data (personal communication M. McLean, November 9, 2015), only 16.1% of the PE 458 principals/school administrators agreed or strongly agreed that PISA/PCAP results were easy to use in making instructional change.

The third inquiry was about other effects of provincial LSAs on the high school. The principal talked about both the grade 11 mathematics and grade 10 literacy assessments. She described the grade 11 mathematics assessment as having a bad influence on students. The principal stated, "It was brutal. The test did not define them." Based on the principal's comments, the LSA of Grade 11 mathematics was extremely difficult for students and the principal did not believe the scores reflected students' abilities. When connecting this comment to those above, the principal was attributing low student achievement to instructional practices that did not benefit from professional development. This finding raises the question about who is responsible for ensuring teachers have adequate skills to deliver the curriculum?

In terms of the grade 10 literacy assessment, the principal commented that it provided her teachers with a starting point to work on students' literacy development. She described

instructional initiatives that examined literacy skills across all curricula. In addition, provincially created rubrics were used to assess students' writing, in particular in the subdomains of writing format and skills, which the principal described as weak. She strongly believed that the literacy assessment data helped improve student learning and teaching practices at the school level.

The next interview question examined implications of using online technology to administer the grade 10 literacy assessment. The principal was in favour of using the technology to assess students' literacy since students had no difficulty with the technology. The principal also noted that the literacy software program would automatically implement some student accommodations such as extra time and script/reading. The principal expressed disappointment at the exclusion of scores from students who received accommodations during the test from the summary of provincial data scores. She further commented that students with special needs were guided to take the accommodations, but their scores were not included in the pool, which made no sense for those students. The other concern related to using online technology was attributed to a miscommunication between the Department, teachers, and students about students' access codes. The principal described this as a disaster. Although the issue with students having their access codes was undoubtedly problematic on the day of administering the LSA and could have probably been avoided, it can be attributed to a dip in implementation that is not likely to re-occur in the upcoming years.

The principal was also asked whether she believed that students' scores on the grade 10 literacy assessment and grade 11 mathematics assessment were good indicators of students' post-secondary attainment. The principal's response was a little hesitant in comparison to

the Achievement Specialist's response. The principal believed that if the assessments were based on the provincial outcomes and students receive high scores, they could possibly be ready for their next step into post-secondary programs. However, she also believed that students' ability, drive for learning, work ethic, attitude for learning, and mental health were other factors that influenced students' post-secondary attainment. Hence she was quite sceptical that students in her school who scored well on the two LSAs also had the non-academic skills that were required for success in post-secondary programs. Unfortunately, the principal did not elaborate on why these skills, which can be argued as necessary for success in high-school programs, were not being fostered in the high school.

The last interview question focused on the challenges that she faced as a high school principal in participating in the provincial and international assessment (PISA) (note: the national LSA occurs in intermediate school). The most challenging aspect of the provincial LSA of grade 11 mathematics was teachers' professional development. She commented that although the Department sent a list of things to focus on, there was no professional development for teachers who were not a part of the curriculum development or marking board. The principal believed "[The teachers need] a lot of professional development on that math test [Grade 11 mathematics test]. They [the teachers] should have professional development on the front dent so that they know how the test is constructed, so that kids could be well prepared, that didn't happen. There was no professional development pulled out, and people come out to intervene. I just thought that they [the teachers] were very much left on their own."

Overall, the principal believed that provincial assessments were very helpful instruments since they informed teachers' instructional practices. She summarized stating,

"I think they give us a hook to hang on and a place to begin." She concluded by noting that provincial assessments were just pieces of the whole picture of learning and teaching and needed to be combined with school data and teacher observations.

Special questions for the representative from the Home and School Federation and the representative from the Bioscience Business Sector. The same questions were posed to the Home and School Federation (HSF) representative and Biosciences Business Sector representative (BBS) since they were both considered external to the direct education of children. The first interview question focused on the skills required by the future workforce to meet the requirements of growing industries, such as aerospace, bioscience, information technology, and renewable energy. The BBS representative explained that strong basic skills in mathematics, science, and English were a foundation from which the future workforce would then be required to specialize and obtain much higher skills. He also added that strong critical thinking and problem solving abilities were important as well.

In comparison, the HSF representative put less emphasis on the learning of content and more emphasis on communication skills. He elaborated that by looking at the data from the Canadian Conference Board, employers were more interested in hands-on problem solving skills and communication skills. He was concerned that "Once you have information to share, you have to communicate. So, communication skills are crucial important. Employers are more interested in communication skills. They are looking for students that are engaging in analyzing, reading newspapers, showing interests in public affairs. They [the students] have an opinion on things and to be able to share and communicate their opinions. That's what I think we should do. I am not sure we are doing it." It is possible that the HSF

representative assumed that students have acquired a certain competency level in the content but he did not directly state that prior to developing communication skills, students must have mastered a content area so they have something to communicate.

The second question posed inquired about whether the information provided from LSAs was a good indicator of students' post-secondary or work/career attainment. The representative from the BBS stated that LSA gave him an idea of students' abilities in mathematics, science and English. He stated, "They [the students] are not performing well all the time in those areas, the amount of training the industrial people have to do is constantly very high and very costly. And it's really [to the] disadvantage of the students, they [the students] don't get the chance to even get into the company. The education is so low that what they are going to do? Our [PE] employment rate is much higher than Canada which is 6.8%, here is about 12% which is not good."

The representative from the HSF believed that the PISA scores did influence some effective changes in PE's education system. However, he mentioned that it was impossible to map the PISA skills since the education system was outcome based rather than skill based. It does not appear the HSF representative understood that the PISA examined broad content areas that students in grade 10 should know. The HSF representative hoped that somebody in the Department of Education would examine all the data from these LSAs and provide effective initiatives and interventions that could enhance student achievement. The BBS representative also indicated that the LSAs were not only for students, but more for teachers. He elaborated, saying "the tests are for the teachers and the policy makers. Teachers constantly ignore the facts of student poor performance. It's totally depending on teachers and their teaching methods to convert student low score to a higher one."

The third question asked of these two participants was what they thought was the main cause of poor PE student achievement in the domains of math and science. The BBS representative said the three main causes were curriculum, quality of teaching staff, and school policies. With respect to the curriculum, he thought that comparisons with other provinces should be done and let teachers know the differences. Secondly, he believed high quality teachers were key to student learning and since the quality of teachers was so pivotal to student learning, teachers' abilities should monitored and measured. In terms of school policies, the issue of missed classes due to snow days was also raised. He suggested a new policy should be put into practice that ensured the number of instructional school days were met and concluded with "if not, they would be doing the same old thing over and over again and expecting different results. That is just ridiculous."

The representative from the HSF also provided three reasons he considered to be the main causes of students' low scores on these LSAs. Firstly, he stated that there was a lack of educational culture. He indicated that many youth in PE do not set education as a priority and neither do their parents. Instead, he believed that sporting events and other outside side activities and events took priority over education, resulting in missed instructional time. The HSF representative explained, "I don't think we yet have a strong education culture on PEI. We see symptoms of that. When in the winter, the PEI absent rate in school on Fridays...because many of them [the students] are hockey…I talked to the principal that I found [at] my daughter's high school, the absent [rate] is a big issue. A lot of that had to do with some students' parents not having a right priority, while sports would take priority over education. That needs to change." The second explanation given for students' low scores on LSAs was students' lack of career aspirations or goals. The HSF representative believed

that students in high schools did not know where they were going after graduate and stated, "The other thing that can affect performance of students in these tests is a real lack of seeing opportunities. Students [middle school and high school] are not really being engaging in the school activities because they don't know where are they going. Many students just don't make decisions. They don't really have a focus or a goal. For education, I think that is an issue." The third reason he explained was due to social media, which he believed distracted students from focusing on education.

The fourth question was posed only to the BBS representative. The question asked about the implications of poor PE student achievement as it related to businesses in general and more specifically, in his own area. He stated "It's very competitive and costly to hire. Every time you hire person, it is [$]300,000 to 500,000 decision. Well, you tell me if somebody is a A student or a C student on PE, and you have a A student from Alberta or Ontario or Quebec, which student are you going to hire? The latter one."

With respect to the field of bioscience, he stated it was a very a competitive field and companies were looking at the best people and because of the poor education system in PE, these students are not going to get jobs in the biosciences. He further added that PE students most likely revert back to manual jobs such as farming or fishing. Alternatively, he said they would leave PE to seek employment elsewhere. He elaborated, "the difficulty is that, if you are looking or building some areas, like Bioscience or IT, you need a large part of your people to be coming from education system here benefiting. It's not very good."

Conclusion

The questionnaire findings told us that there were still quite a few participants (around 30%) who had no idea about PE student performance on LSAs. There was not much

difference between parents and nonparents regarding their knowledge of LSAs. For those participants who knew about LSAs, they were more likely to know about LSAs function in education accountability in a broad way, while holding a neutral or slight negative attitude of using LSAs as instruments indicating school quality and student success. Many, especially parents, also had a surprisingly negative view on the usefulness of LSA as a means of providing information about the health of the educational system. In contrast, opposite views were held by the interview participants, possibly because of their work and personal experiences with LSA. The link between LSA scores and the prosperity of the local economy was not perceived by participants, which revealed less awareness of the broader functions of LSAs over their function of providing information about the wellbeing of the education system.

When comparing perceptions towards LSAs with the three grouping variables, parental status, educational attainment, and cultural affiliation, there was no significant differences between either of the groups.

The interview data indicated that all four participants were quite familiar with the current implementation of LSAs in PE. In terms of the participants perceptions related to the accountability function of LSAs, all of the interview participants admitted that, theoretically, LSAs are good indicators of the education system's health. However, they found it difficult to see the PE education system actually using the LSA data to fulfill functions such as gatekeeping and monitoring student progress in learning, primarily because of a lack of professional development for teachers related to LSAs. Since the grade 10 literacy and grade 11 mathematics assessments are just in the piloting stage, the interview

participants indicated that none of the provincial assessments played the role of gatekeeping before these two newly introduced LSAs.

In terms of the latter two functions of LSAs, diagnosing and monitoring student achievement, although the high school principal positively demonstrated using provincial assessment data to diagnose students' current state of learning for further teaching strategies, it is still hard to conclude that all high school teachers and principals would share the same positive value of LSA and make actual change in their classroom instruction as a result. No feedback was provided by any of the interview participants that school administrators or teachers would use LSA data to monitor student achievement. The need for more professional development for teachers to understand and make good use of data was raised by the high school principal, which she considered as the main barrier to good utilization of LSAs on PE. This feedback was well aligned with the literature review, where lack of teacher professional development seems a common problem hindering the fulfillment of implementing LSAs in any meaningful way.

PE students' low position on national and international LSAs triggered the interview participants' concern of the quality of the education system to ensure students' mastery of basic knowledge in literacy and numeracy as well as advanced soft skills such as critical thinking, communication, and problem-solving. However, according to the interview participants, this basic knowledge was essential for students to compete in future career development. Lack of fundamental knowledge in literacy, numeracy creates barriers for students to get employed on PE as well as other places, which is not benefiting the economic development of the province either.

These four interviews provided a more in-depth understanding of current conditions, issues, and economic influence of LSAs as practiced presently in PE, together with some thoughts on future developments by people in affected fields of education and business. The quantitative data showed that approximately 40% of participants knew the true state of LSA results in PE, and believed that the LSA data provided good indicators of education system accountability. However, they did not believe it can also be an important indicator of teaching quality. This lack of teaching accountability is also indicated by the disregard shown by the Teachers' Federation's unwillingness to participate in this study. Together with the quantitative data, the interviews thus provided depth to the broader understanding of large-scale assessment in Prince Edward Island.

CHAPTER SIX

Discussion

It is widely believed that large-scale assessments (LSAs) are designed and implemented as instruments for measuring students' achievement in educational programs (Nagy, 2000). Thus they provide useful information about the accountability of the education system and its ability to produce a qualified labour force to fill future careers that sustain and enhance economic development. International and national organizations such as the OECD, CMEC, Conference Board of Canada, and regional institutions of market studies (e.g., Atlantic Institute for Market Studies) have been documenting the relationship between LSAs and economic prosperity. It is widely believed that a knowledgeable and skilled workforce enhanced with critical thinking skills and the ability to innovate would largely benefit economic growth and improve the overall life-style in society (OECD, 2014b; The Conference Board of Canada, 2015b). Hence, objective measurements of students' learning outcomes were of great importance to stake holders (i.e., policy makers, school administrators, communities, parents, and students) as they would monitor learning progress thus LSAs were pivotal in the advancement of the global economy.

This chapter synthesizes current literature documenting the significance that LSAs play in the global economy with the findings presented in the previous chapter to respond to the two research questions posed in this thesis. The response to the first research question begins with a review of the hypothesis and is followed by a discussion on what participants know including a focus on what parents know about students' LSA performance. Since approximately 65% of the participants were informed about students' mathematics performance, the next section focuses on possible explanations why knowledge about

mathematics performance was higher than in other domains. Other factors influencing participants' LSA knowledge focused on educational attainment and cultural affiliation. The discussion for the first research question concludes with a reflection on issues related to accessing LSA information.

The discussion on the second research question focuses on participants' perceptions towards the utility of LSAs as instruments of accountability as measured using a 12-item scale. This section begins with an examination of responses to the individual items specifically related to indicators of accountability including educational system, school, and teacher accountability. The outcome of the factor analysis is presented next followed by findings from the analysis of variance where mean scores from the scale measuring perceptions towards LSAs is compared with the three grouping variables: parental status, educational attainment, and cultural affiliation.

The tail end of the chapter discusses the limitations of the study focusing on the shortcomings in the questionnaire and small sample size of participants who belonged to the Aboriginal culture and participants whose educational attainment was below Grade 12. Lastly, two implications and two recommendations of the study are summarized.

Research Question 1: What are the public's understandings of PE students' performance on large-scale assessments?

Even though LSAs occur at the school level, the impact of students' performance extends beyond the school community. Nagy's (2000) accountability perspective of LSA highlights stakeholders in education to whom educators are responsible. These stakeholders include parents but also other stakeholders such as businesses because they are also affected by quality of the education system (OECD, 2014a). It can also be argued that educators are

responsible to health care providers given the impact education has on ones' health. It is well documented that less educated people place greater demands on the health care system (Baum, Ma, & Payea, 2013). Hence, those responsible for educating children are not only accountable to parents but also businesses, health care providers, and other social service providers who are collectively described as the public.

Hence, the public's opinions on LSAs matter because the public are the taxpayers who fund the education system that prepares children to fulfill the employment demands of the 21st century. Thus, it is the public who are responsible for holding educators accountable. Prior to examining the public's opinions about LSAs it is prudent to first explore what the public knows about LSAs. Further, it is important to examine whether some groups are more knowledgeable than others. For instance, it is likely that parents are more knowledgeable about LSAs given the direct impact LSAs may have on their children (Goodall & Vorhaus, 2010). It is also possible that people who have higher credentials are more informed about the impact of LSA on the economy or health hence they may be more attuned to LSA results (Baum, Ma, & Payea, 2013). Lastly, people's cultural affiliation may influence their knowledge about LSAs as the OECD (2012) reported that the PISA data allowed the public to compare the outcomes of compulsory education which indicates the quality of the education system, thus may provide incentive for people moving to other countries.

Hypothesis. Given PE's low-stakes nature of the current LSA program combined with the relatively recent implementation of LSA in the province and the weak performance of PE students on both national and international LSAs (CMEC, 2001[PISA]; 2004 [PISA]; 2007 [PISA]; 2008 [PCAP]; 2010 [PISA]; 2011 [PCAP]; 2013 [PISA]) (except for the most

recent PCAP in 2013), it was hypothesized that a large portion of the PE public would not be well informed about PE's students' performance on LSAs. This thought is rationalized further highlighting the *consistently* low performance by PE students, in more than one domain over the 18 years that LSAs have been administered at the national and international levels. It was this pattern of performance that influenced a belief about the public's lack of knowledge regarding students' LSA scores. Although PE students made improvements on the last PCAP in 2013, it is unknown whether participants would be aware of this improvement, particularly, if the previous and consistently low performance was etched in their minds.

Participants' LSA knowledge. To explore participants' understanding of LSAs, they were presented with six items exploring their general knowledge about student performance on LSA. The scale for the first four items (i.e., items 10 a – d), surveying knowledge about PE student achievement in reading, mathematics, science, and problem-solving domains, included a response option to indicate they did not know anything about students' performance in these areas. This response option nicely captured participants' absence of awareness about LSA where a range of 23.7% to 36.1% of the participants admitted not knowing anything about LSA in any of the four domains.

Investigating further, response patterns to the "don't know" response in Item 10 from only parents were examined. In this case, it was expected that parents with children currently in the schooling system would be more knowledgeable about students' LSA performance than non-parents given that their children would be experiencing LSAs periodically throughout their schooling. Recalling, the total number of participants who indicated they had children currently in the public school was 256. Of these parents, 23.1%

to 36.1% admitted they did not know anything about students' achievement on LSA (note: the range was very similar to the entire population). More parents reported not knowing about students' performance in problem-solving (36.1%) and the fewest parents reported not knowing about students' mathematics performance (23.1%).

When participants were asked to compare PE's students' achievement in relation to other provinces on the national (Item 15a) and international (15b) LSAs, the majority of participants were not aware that PE students' had made gains in achievement on the most recent PCAP in 2013 in comparison to their achievement on the three preceding cycles of assessment. On this item, 78.6% of participants indicated that PE students scored below the national average when in fact they were about average. In contrast, the majority of participants were knowledgeable about students' performance on the PISA where 81.9% of participants indicated students scored below the provincial average.

When comparing response patterns between parents and non-parents, approximately the same percentage of parents and non-parents had accurate understandings of PE's students' performance on the PCAP and PISA. The important finding in this item was the small percentage of parents and non-parents who were not current about PE's recent improvement on the 2013 PCAP. Further, there was a large percentage of the PE population of both parents (82.9%) and non-parents (81.1%) who knew that PE students' did not score well on the international LSA.

Although the majority of participants in this study knew that PE students did not score well on the PISA, there was a lot of misinformation about students' PCAP achievement. This finding about participants' PCAP knowledge aligned with previously noted research about parents' perceptions and beliefs about LSAs in Ontario where it was reported that

approximately 40% were not aware of the data provided by their provincial testing body (Mu & Childs, 2005). These findings also echoed the work of Cirtwill, Clifton, and D'Orsay (2002) who reported that a 'large proportion' (this descriptor was not quantified) of parents did not know anything about their child's performance at school, or were ill informed about students' achievement due to insufficient data. It is important to note that Cirtwill et al.'s study occurred before 2002, when PE had not yet started provincial assessment but had engaged in the SAIP (national assessment during that time period) and the PISA (international assessment). At the time of this thesis, PE had been involved in provincial assessment for nine years, which was thought to have raised some awareness about LSAs in general. Hence this thesis not only supports Cirtwill et al.'s research but also provides a larger and more recent sample in the PE context to substantiate it. As well, this study quantifies that a large portion of the population who do not know about students' achievement on the PCAP is approximately 80% of the population but caution is also exercised because only 18% of the population were not familiar with students' PISA performance. The large number of participants who were informed about students' achievement on the PISA is contrary to what was hypothesized.

Participants' knowledge about mathematics performance and possible factors influenced their knowledge. Recalling, there was a small percentage of participants who were not knowledgeable about students' mathematics performance in comparison to the three other domains. While factors influencing participants' knowledge may be indirectly attributed to PE's recent LSA initiatives and the publication of results, it is more likely that knowledge was influenced by recent media publications which called attention to PE's poor student performance; especially in the area of mathematics achievement. Recalling, 65% of

participants indicated that PE was at or very near the bottom on the LSA of mathematics (Item 10b). This large percentage of participants who had an accurate understanding of PE's student achievement in mathematics, may have been due to recent and numerous media reports on the radio (i.e., Island Morning), in the local newspaper (i.e., *Guardian*), and in the national newspaper (i.e., *Globe and Mail*), which was acknowledged by two of the interview participants: Home and School Federation and Bioscience Business Sector representatives.

Media coverage of PE's student weakness in mathematics was raised in a CBC radio interview with Dr. Tess Miller from the University of Prince Edward Island, hosted by Matt Rainnie, where Dr. Miller shared student achievement trends in mathematics on national and international assessments. This radio interview was followed by a public address billed as the State of Mathematics Education in PE in which over 200 people from all sectors of PE attended the address presented by Dr. Miller. *The Guardian,* followed up with an article, called "Poor math scores in PEI have reached a crisis point." In this article the author, drew from Dr. Miller's presentation highlighting her concern that students' poor mathematics performance would influence students' success in high school and post- secondary achievement which would have long-term consequences that might present barriers hindering students' career goals (Wright, 2014a). A second article in *The Guardian* that also highlighted PE students' poor performance in mathematics was called "Poor math skills just don't add-up, opposition leader says." (Wright, 2014b). During the 2015 provincial election campaign, issues related to LSA were debated and reported in *The Guardian*, which were also likely to influence participants' LSA knowledge (*The Guardian*, April 12, 2015). Lastly, another media release in a national newspaper that may have influenced participants' knowledge about PE's student achievement was an article "Math wars: The

division over how to improve test scores" from *The Globe and Mail* (Carson, 2014), which pictorially showed PE as the lowest ranking province on the recent PISA in mathematics. Although information communicated through radio and newspapers do not reach all members of a community, the information is likely to be heard by the majority of people and in the context of this study, information about students' LSA achievement in mathematics may have been heard by 65% of the sample. Further, since this information is recent, within the past couple of years, may explain why there has been no changes in LSA scores, other than the most recent PCAP (2013), on which PE student scores increased by 32 points from 460 in mathematics.

Participants' knowledge about PE student performance in other subjects. There have been less media comments or reports that have specifically focused on how PE students scored in the other three domains of reading, science, and problem-solving. For those participants who were knowledgeable about student performance on these three domains, they may have obtained the information from the national and international websites such as the CMEC or the OECD. Also, along with the focus on mathematics performance, some of the local media reports had presented PE rankings and the scores in all four domains. It is important to note that the science and problem-solving domains are not assessed on the provincial assessments, thus there has been no provincial data from which to draw attention to these domains. This may explain the large percentage of participants who were not informed about student achievement in science and problem solving in comparison to mathematics or reading.

Other Differences in Participants' LSA Knowledge

Educational attainment. Differences in participants' LSA knowledge were examined based on educational attainment. In terms of participants' responses based on educational attainment, there were slightly more college/apprenticeship participants who knew about students' PCAP achievement than the other educational levels. In contrast, there was a greater percentage (84.8%) of participants' holding graduate certification that accurately knew that PE students scored poorly on the PISA. The least informed group was participants' holding grade 12 or less certification (73.5%).

Given PE's history of poor LSA results at the national and international levels, it is possible that their historically poor performance has shadowed students' improvement on the recent PCAP in 2013 where participants are not aware of recent progress. Hence, if we focused on the PISA we can conclude that the participants in this study are knowledgeable about students' poor achievement. In addition this finding is aligned with the trend showing the impact of higher education on society in as much as those who are more educated, are more likely to be employed, have healthy lifestyles, and have the potential to make society a better place (Baum, Ma, & Payea, 2013). In extending this thinking, it is also likely that those who are more educated would be aware of the trends in education. In the case of the participants in this study, there was a smaller percentage of participants holding a grade 12 or less who were knowledgeable about the students' PISA scores. In comparison, there was a greater percentage of participants holding undergraduate and graduate degrees who were knowledgeable about students' PISA scores.

Cultural affiliation. In examining participants' LSA knowledge as reported in item 15, grouped by cultural affiliations, it was hypothesized that a greater percentage of

Newcomers would be more knowledgeable about students' LSA performance because the OECD reported that people considering moving to other countries reviewed the results (OECD, 2012). In this study, this was not the case. When grouping participants based on their cultural affiliation, there was a small percentage (19.2% - 25.0%) from each culture who accurately knew about students' PCAP performance and a large percentage (77.6% - 83.3%) of participants who accurately knew PE about students' PISA performance. Each cultural group responded within 3% of the overall average for both the PCAP and PISA.

Public Access to LSA Information

The topic of access to LSA information affects participants' knowledge about students' LSA performance. If the data are not accessible, then participants are not going to know how PE students perform. The topic of accessibility to LSA information was presented in the literature (Livingstone, Hart, & Davie, 2001; Wu & Childs, 2005) and was also raised during the interviews. Previous research reported that Ministries and Departments of Education, School Boards, parent groups, and education related organizations (e.g., EQAO, Fraser Institution, C.D Howe Institute) were the main sources of information from which the public is able to obtain information about LSAs (Wu & Childs, 2005). Wu and Childs (2005) reported that parents tended to underutilize the detailed information provided by the provincial testing agency (i.e., Education Quality and Accountability Office [EQAO]) and that only 13.5% of the parents visited EQAOwebsite. In this present study, the questionnaire did not survey where participants obtained their LSA information; however, the interview participants commented on where they received theirs. The achievement specialist and the high school principal, who both worked for the Department of Education indicated they received their information directly from the Department of Education which publishes

student achievement information. However, both of these participants received more detailed information related to test items and students' performance grouped by teachers and schools. With respect to the other two participants, the Home and School representative received individual, school, and provincial results from his daughter's teacher. He also tracked national and international assessment results from the CMEC website for his own personal interest. The representative from the Bioscience Business Sector indicated that he gathered his information from websites as well as the media.

The PE's Department of Education website serves as a public source of information about LSAs. However, after thoroughly examining the website as well as the websites in other jurisdictions, the PE's Department of Education's website releases little information about their LSAs and analysis of students' performance. The PE Department of Education does provide sample assessment instruments with implementation guidelines for teachers or school administrators as well as each schools' overall performance from the most recent assessment but not the performance from prior years. Less information about students' LSA performance was found on the two school boards' websites.

A report documenting PE students' national and international performance on the provincial website was simply adapted from national and international assessment information reported on the CMEC's website without any additional analysis specifically in PE context. The absence of provincial LSA information on PE's government website could be the reason why approximately 50% of participants were not knowledgeable about PE's student performance on LSAs. Further, the absence of comparative data showing, for example, the large number of PE students scoring in the lowest level of achievement and the few number of students scoring in the highest level of achievement on the provincial

assessment could also be contributing to participants' lack of knowledge. Showing the distribution of students' performance by levels drawn from provincial, national, and international data may possibly be what is needed to raise more awareness in the PE population about LSA scores which may act as a catalyst to advocate for change in the education system. Without a concerted effort by the public or advocates of education (e.g., Education 20/20) to inform the general public through various news media (Sundgren, 2005), there is little opportunity for people to gain knowledge, obtain information about LSAs, or understand the impact of students' LSA achievement on the wellbeing of the community. A third party similar to the Educational Quality and Accountability Office in Ontario who are responsible for sharing LSA data as well as monitoring the usage of LSA data and publishing research based on the data, may be what is needed in PE to advocate for the betterment of education. Without access to all data displayed in readable and informative contexts, the accountability function of LSAs is in jeopardy.

Personal reflections on gathering provincial LSA data on PE students' LSA achievement. Prior to concluding this section responding to the public's understanding of LSA, it is fitting to reflect on my journey to gather LSA data and information from PE's provincial assessment program as it speaks to the lack of transparency in a publicly funded program that is supposed to hold educators accountable. The first information source I looked at was the PE Department of Education website. On this website, some basic information about provincial assessments were provided, such as a guide for parents, administrators' guide, and sample paper for each of the Grade 3, 6, and 9 assessments. From 2014, provincial assessment results can be checked by individual schools on the website for the most recent assessment. This school-based report shows the number of

students who wrote the test, the average score based on school level, and the percentage of students who met the standard on this current or most recently published data. However, there is no report based on student performance at the provincial, School Board, or school level in terms of average scores and the percentage of students at each level of achievement based on the provincial standards. Furthermore, no descriptions of the rubrics used to measure student achievement were presented to describe the standards at each level of achievement. Hence, when the Department of Education reports a percentage of students who met the standard it has no meaning since the general public (including myself) have no understanding of what the standard was.

In order to understand how students performed from year to year at the key stages in the education system, I searched the provincial website for the annual reports released by the Department of Education of PE. Buried in the reports, was a section listing the provincial assessment results from year to year at Grades 3, 6, and 9 in the domains of Language Arts and Mathematics. The results presented the percentage of students who were a) approaching the standard, and b) met the standard, from year to year on the provincial scale. This presentation of student performance from year to year is not available on the provincial LSA website. Missing from the annual report was a breakdown of the number of students by levels of achievement as reported by other jurisdictions (e.g., Ontario, British Columbia, Quebec).

To obtain a more thorough understanding of students' performance on provincial assessments, I corresponded through email with a staff member from the Instructional Development and Achievement Office of Department of Education asking for detailed reports. The documents I received on the Language Arts and Mathematics reports were

separated by school boards and school districts, from 2007 to 2013. For Language Arts, it provided the percentage of students who actually met the standards, moreover, it also showed the percentage of students who met the standard in both subdomains (i.e., personal expressive and transactional) on the writing assessment.

Unfortunately, this information cannot be obtained from public sources, but is important to thoroughly understand students' achievement in PE. This information revealed a large difference in the number of students performing in each of the three levels of achievement (i.e., experienced difficulty, approached standards, and met standards) on the annual provincial assessments over a six-year cycle. To contrast the depth of information between the annual report and the information from the Instructional Development & Achievement Office, take for example, the Grade 3 mathematics assessment. In the 2010-2011 school year report, it revealed that 83% of the students were at or approaching the standard, however, the report from the Instructional Development & Achievement Office of the Department of Education, revealed that only 72% of the students actually met the standard and 11% of the students were reported as approaching the standard. The Department of Education's combined method of reporting is misleading. Further, the criteria for students who were classified as approaching the standard is unknown. This reporting practice is likely to create a false impression for the public who, if they have gathered some information about students' provincial achievement, are led to believe that more than 80% of students have met the standard, which is not true.

The other difficulty in the manner of reporting students' provincial achievement is that both reports from the website of the Department of Education and the Instructional Development and Achievement Office do not include any detailed descriptions related to the

provincial assessment standards. Hence the public or researchers like myself, have no idea what a student who has met the standard can actually do! This aspect of PE's LSA reporting practice is quite different from other provinces (e.g., Ontario, British Columbia, Quebec) as well as national and international assessments.

To aid the public in accurately understanding students' LSA performance in the province, all data must firstly be made available. Secondly, the data must be presented by achievement levels.

Research question 2: What are the public's perceptions towards the utility of large-scales assessments as instrument of accountability and how do perceptions differ based on participants' parental status, educational attainment, and cultural affiliation?

The following section focuses on the public's perceptions about LSAs' in terms of their perceived importance of the LSAs functions as mentioned earlier in the studies of Nagy (2000) and Klinger, DeLuca, and Miller (2008). If the public perceives LSAs as important instruments that hold the education system accountable, as espoused by Nagy (2000) in his theoretical assumptions underpinning LSAs, the public, especially the parents, would be more in favour of using LSA data as an indicator of school quality, teaching performance, and students' academic competencies, as well as students' prospects for future careers and quality of life.

Introduction to 3 grouping variables. Further, the CMEC and OECD's studies, as highlighted in the contextual chapter, revealed that family background, which mainly refers to family structure (i.e., single parent or two parents), and whether the student is from an immigrant or non-immigrant family(CMEC, 2001; CMEC, 2007; OECD, 2003) and parents' educational attainment (CMEC, 2004; OECD, 2003) influences students' academic

performance. OECD also claimed that student who had more cultural possessions such as classic literature and work of arts scored higher than average. In addition, communicating social and cultural issues also helped in improving student performance on PISA, especially in reading domain. Therefore, in this study, it was hypothesized that participants' parental status, educational background, and cultural affiliations would influence participants' perceptions towards LSAs. More specifically, it is assumed that participants who have children in the public education system would respond more favourably towards the accountability purpose of LSAs than non-parents. It is also reasonable to assume that participants who have higher education would have more positive perceptions towards LSAs as they may have a better understanding of LSA purposes. As for the cultural affiliations, some cultures such as China (Cronenweth, 2012) have established a long history with LSAs where they have become common place and even valued instruments that determine who is accepted into the best institutions. Although, participants' individual culture was not documented, the majority of immigrants in PE are from Asian countries (CBC News, 2014) hence, it is possible that Newcomers may hold positive perceptions towards LSAs based on their prior LSA experiences.

The following sections synthesize key findings from the survey and interviews to explore responses to individual items that explore issues of accountability. Following is a discussion on the comparison of mean scores based on the three grouping variables: parental status, educational attainment, and cultural affiliation.

Accountability indicators of LSA. As previously noted, LSA programs serve as instruments for policymakers to make judgments about the overall quality of the education system and the quality of teaching (Nagy, 2000; Klinger & Rogers, 2011). Items 13a and

14c from the scale specifically focus on the educational system accountability function of LSAs. Item 13a revealed that 54.9% (*M*=2.57, *SD*=1.58) of the participants responded with a 1 or 2 believing that LSAs revealed a lot about whether the education system was working. In order for the accountability function of LSAs to be working it seems reasonable that the public should believe that LSAs are indicators of the effectiveness of the education system as espoused by Nagy (2000) and Klinger and Rogers (2011). The question that remains unanswered is how many people need to believe that LSAs are indicators of the effectiveness of the education system for the accountability function to be effective? Although we do not have the answer to this question, we move onto Item 14c to add more insight. Item 14c examined perceptions of accountability specifically focusing on how the PISA was perceived as an indicator of the quality of schools. For this item, only 37.8% (*M*=3.04, *SD*=1.651) of participants responded at the positive end of the scale with a 1 or 2 indicating they believed that the PISA provided information about the quality of schools. Based on participants' responses from these two items, we can conclude that barely half of the population believes LSAs are instruments of system accountability. Although, there is no other benchmark to contrast this finding with, it is unlikely that this is a sufficient mass of the population to hold the education system accountable.

Nagy (2000) also described the utility of LSAs as instruments for teacher accountability. He pointed out that although teachers always engaged in instructional diagnosis, they lacked the ability to utilize external assessment (i.e., LSA) data to inform themselves about student achievement to further improve their teaching practices and enhance their professional development, which subsequently compromised the accountability function. As Volante (2007) noted, educational accountability is primarily a relationship

between taxpayers, elected officials, and teachers. Hence it is important to know how the public perceives LSAs as an indicator of teacher quality. Items 13f and 14e focused on the utility of LSAs as instruments for teacher accountability. Recalling, Item 13f, asked participants whether they believed LSAs provided information about whether teachers were doing their jobs. The responses to this item were fairly well distributed between believers (responded with a 1 or 2), middle group (responded with a 3 or 4), and non-believers (responded with 5 or 6): 39.2%, 37.2%, and 23.7%, respectively (M=3.13, SD=1.671). An examination of responses to item 14e, which examined the extent to which participants believed that the PISA provided information about the quality of teachers, revealed a similar distribution of responses grouped by believers (30.2%), middle group (33.1%), non-believers (23.4%) (M=3.31, SD=1.076). Based on the responses to these two items, the majority of participants in this study do not believe that LSAs are an indicator of teacher quality. This finding echoes the work van Barneveld (2008), Volante and Cherubini (2010), and Wu (2010) who argued that it is not necessary to link the results from LSAs with teacher performance. If this is the case, then Nagy's accountability framework for LSAs that call for teacher accountability may be flawed. More so, if the majority of the public does not believe that LSAs are an indicator of teacher quality then the accountability function of LSAs is not working in this province.

Since there has been a great deal of research indicating that teaching quality is the most influential variable in student success (Darling-Hammond, 2000; Harris & Sass, 2011; Strong, Ward, Tucker, & Hindman, 2008; Rowe, 2003; Huitt, Huitt, Monetti,& Hummel, 2009), the perceptions of teachers and retired teachers were selected from the data set and examined. Not surprisingly, both teachers and retired teachers were not in favour of using

LSAs as an indication of teacher quality (Item 13f: Teachers: *M*=4.14, *SD*=1.582; Retired Teachers: *M*=4.33, *SD*=1.966) (Item 14e: Teachers: *M*=4.70, *SD*=1.636; Retired Teachers: *M*=5.00, *SD*=1.673).

The perception from interviewees about LSAs as instruments of teaching accountability was different than what was revealed in the survey. The representatives from the Bioscience Business Sector and the high school principal shared a common view. The Bioscience Business Sector representative noted, if teachers were responsible for students' learning outcomes, then based on students' LSA performance, he questioned teachers' capacity to help students' improve as measured on LSAs. This interview participant expressed major concerns about PE teachers' active role in implementing LSAs and their ability to interpret the data in order to guide their own classroom instruction, therefore enhancing students' performance. These perceptions of teacher accountability were also shared by the high school principal. The high school principal added by reflecting on her recent experience administering the pilot version of the new Grade 11 mathematics LSA. She attributed low student achievement on that provincial LSA to teachers' inability to prepare their students because they did not know what to expect in terms of content and depth of assessment. Since the provincial LSA of Grade11 mathematics was not based on a new set of curriculum documents, one would expect that teachers should have been quite familiar with the content and depth of assessment. However, this is possibly what plagues PE students' performance; a shortfall in the content coverage and depth of inquiry, which can be due to a number of problems. Such problems include high student absenteeism, students who are unprepared to learn the prescribed curriculum. These problems do not exclude, teachers' lack of pedagogical or content skills. Although it seems feasible that an

examination of teaching quality is warranted, the probability of conducting such research is difficult given teachers' and schools' reluctance to participate in research (Alibali & Nathan, 2010; Samaroo, Dahya, & Alidina, 2013). As noted at the beginning of this thesis, attempts were made to solicit teachers to participate in the study but after several attempts, one teacher rejected the invitation to participate and a second teacher neglected to respond to several emails. Such a reluctance to participate in LSA research may be due to the influence of the PE Teachers' Union who have vocally denounced LSA (Wright, 2013; Zwaagstra, 2014). Another explanation attributed to poor student performance was teachers' ability or inability to understand and appropriately use LSA data to guide their classroom instructions. As noted previously, teacher professional development has been well advocated (Volante, 2004; Volante & Cherubini, 2010; Klinger, Volante, & Deluca, 2012). It is important to note that although the principal advocated for teacher professional development, she advocated for it in advance of the LSA to help them better align their teaching with what is being assessed rather than professional development to learn how to interpret and use LSA data.

Perceptions of LSA. Participants' perceptions towards LSA were captured using a 12 item scale that posed questions related to the amount of information provided by LSAs (Items 13 a - f) and an additional six items exploring the extent to which participants agreed to statements about LSAs and the quality of students, teachers, and schools, use of LSAs to compare countries, and who reads LSA data (Items 14 a – f). Based on this scale, participants held neutral perceptions towards LSA (M=3.39, SD=1.38). Exploring further, perceptions were grouped by parental status, educational attainment, and cultural affiliation.

Grouping by parental status. When focusing on the differences in responses from parents versus non-parents, there was no statistically significant difference between these two

groups in terms of the perceptions towards LSA. Overall, parents with children currently in the school system were slightly more positive towards LSAs than non-parents (Parents: *M*=3.34, *SD*=1.33; Non-parents: *M*=3.47, *SD*=1.45). However, both parents and non-parents responded similarly and positive when asked whether they agreed that PE's student achievement was compared to other countries (Item 14d). On a second item (Item 13e), both groups held negative perceptions about the extent to which LSAs provided information about the success of businesses on PE (Parents: *M*=4.29, *SD*=1.59; Non-parents: *M*=4.41, *SD*=1.72).

Although interviews were not solicited from parents, the representative from the Home and School Federation shared some perceptions about LSAs from his position as a father with a child currently in the school system. He described how his daughter's teacher incorporated part of his daughter's provincial LSA score with the teacher-derived score to generate a final score in a course and this practice was not consistent throughout the province. This scenario raises the issue of teacher accountability in that teachers may not be following the Department of Education's guideline for using LSA scores to determine students' final grades which is aligned with a finding in Miller's (2013) study. As a result of this inconsistency in which students' grades are determined, this parent, has reservations about the utility of LSAs in determining grades.

Grouping by educational attainment. According to the CMEC (2007), parents' educational attainment is a factor that has the potential to influence students' performance on LSAs (CMEC, 2007). Hence, it is reasonable to expect that higher educated people may hold more positive perceptions towards LSAs because they may understand the relationships between LSAs and economic prosperity, health, and children's future career goals. The

ANOVA comparing mean scores for each of the five different educational levels (i.e., grade 12 or below, college/apprenticeship, undergraduate, graduate (Masters), graduate (PhD), did not reveal a significant difference.

Since there was similarity in response patterns from parents and non-parents based on the relationship between LSAs and economic prosperity (Item 13e) where both parents and non-parents did not believe in the relationship, it is not surprising to also see that participants holding a undergraduate education or higher, were skeptical of this relationship (average mean for these four groups: *M*=4.21, *SD*=1.65). In contrast, participants who held a college credential were slightly more positive (*M*=3.62, *SD*=1.64). Once again, responses from the college/apprenticeship group have not been aligned with response patterns from educational attainment groups.

Cultural affiliations. As for participants' cultural affiliations, there was no statistically significant difference in response patterns based on the scale examining perceptions towards LSAs. Newcomers were slightly more positive (*M*=3.32, *SD*=1.33) than participants who identified as belonging to the French (*M*=3.37, *SD*=1.58), Aboriginal (*M*=3.51, *SD* = 1.52), or local community (*M*=3.43, *SD*=1.38) groups. When examining response patterns to individual items based on cultural affiliation two anomalies stood out. Firstly, each cultural group, including the local community, did not believe that LSAs provided information about whether businesses were successful on the island (Item 13e). This counters the OECD (2014a) report showing the connection between student performance as measured on LSAs and economic growth of which the success of businesses is an indicator. This finding also counters what the Biosciences Business Sector representative discussed in his interview about the difficulty in hiring skilled people from PE

to fulfill high-tech positions. Although this finding is based on one item and lacks the power held in a scale (i.e., a large group of items measuring a construct) it does identify an area for further research as well as an area where educators and businesses need to promote to communicate the connection between education and the economy.

The second anomaly was the large percentage (61.8%) of Newcomers who strongly believed (responded with a 1 or 2 at the positive end of the scale) that LSAs provided information about the quality of schools. This contrasts with the French (40.6%), Aboriginal (35.7%), and local communities (36.7%) who were much less positive about the connection between LSAs and the quality of schools. Again, although not a powerful analysis, this finding warrants an opportunity to educate many people about the connection between LSAs and the quality of schools. It is likely that Newcomers are able to contrast their schooling experiences in other jurisdictions or countries with their experience in PE schools. Although participants' previous homeland was not surveyed, it can be assumed that the large number of immigrants (i.e., Newcomers) to PE in the past five years were of Chinese origin (CBC, March 4, 2014), who have a well developed history of LSAs stretching back over more than a millennium and have scored much (PISA 2012 overall score for Shanghai China: 613 in Mathematics, 570 in Reading, 580 in Science; PE: 479 in Mathematics, 490 in Reading, 490 in Science) higher on the PISA in comparison to PE students.

Moving Schools. If the accountability function of LSAs is working then we expect people would do something to advocate for change thereby holding those responsible for educating children accountable. Prior studies have shown that parents would look at academic achievement and teacher quality as the most important factors when they choose a

school for their child (Armor and Peiser, 1997; Greene et al., 1998; Kleitz et al. 2000; Vanourek et al. 1998). In Ontario, the Fraser Institution helps parents with this decision by providing a ranking of schools based on a school's demographic information and students' LSA performance. Jason Clemens, Fraser Institution executive vice-president, claimed: "When parents are empowered to choose their schools for their kids and we force schools to compete with one another, we get better school performance." (CBC News, February 27, 2014)

To explore whether participants would do something about poor LSA scores, participants were presented with a scenario that asked them if they had the opportunity to move their child to a better performing school, based on LSA scores, would they? Responses from the entire sample were distributed equally between the two options where half of the participants would move to another school and the other half would not move. Response patterns were examined further by grouping participants by parental status, educational attainment, and cultural affiliation. The parent groups were evenly divided with a 50 – 50 split. In terms of educational attainment, 68% of PhD holders would move their child to a better performing school in comparison to only 60% of Grade 12, 54% of college/apprenticeship, 46% of undergraduates, and 40% of Master's certificate holders would move schools. When grouping by cultural affiliation, there was a greater percentage of participants from the Newcomers (61%) and Aboriginal (61%) groups would move schools in comparison to the French (41%), and local community (46%).

Given that there is a large percentage of participants who would not move to a better performing school, more research is needed to explore factors influencing their decision to keep their in a poor performing school. In hindsight, the scenario in which this item was

situated could have been improved by eliminating any common reasons for not moving to a better performing school (e.g., transportation, don't want to move to a different neighbourhood, etc.)

Conclusion

The aim of this study was to investigate the extent to which participants' were knowledgeable about LSAs and explore public perceptions towards the utility of LSAs as instruments of accountability in PE's education system. With half of the sample population were uninformed about students' performance on LSAs, the hypothesis seems to be true.

When comparing participants' knowledge about students' performance on LSAs with three grouping variables: parental status, educational attainment, and cultural affiliation, it was found that parents' knowledge was slightly higher than those of non-parents (10%), which aligned with previous studies. With respect to the educational attainment, there were more uninformed college and apprenticeship participants compared to participants with high school certificates in each of reading, mathematics, and problem-solving; which did not support the relationship where an increase in education credentials paralleled an increase in LSA knowledge. When referring to cultural affiliation, Newcomers' lack of knowledge about students' performance rejected the previous hypothesis. This finding may be attributed to the language barrier preventing Newcomers from receiving information about PE students' performance on LSAs.

In summarizing other key points stemming from the research question it is important to recall the large percentage (65%) of participants who had an accurate understanding of PE student achievement in mathematics. This finding may have been due to recent and numerous media reports from the newspaper or on the radio. Another key point discussed was the personal challenges experienced in gathering LSA data which is a testament to the lack of transparency in the reporting of provincial LSA data.

The second research question explored public perceptions about LSA's function in education and teacher accountability. There was a large percentage of participants who reported that LSAs in general provided information about the education system. However, when participants were asked if the PISA provided information about the education system, a approximately 20% less agreed that the PISA provided information about the education system. Hence, barely half of the sample perceived LSAs were instruments of educational accountability.

Discussions related to teachers accountability, drew from both the questionnaire and the interview and revealed that the majority of participants in this study did not perceive that LSAs are an indicator of teacher quality. Thus, although the relationship of teachers quality and student achievement is well documented in the literature, if the majority of the public does not believe that LSAs are an indicator of teacher quality then the accountability function of LSAs is not working in this province.

Public perception of LSA was examined through a 12 item combined scale, further grouped by the three variables. Both parents and non-parents responded similarly and positively about the PISA's indication of comparing PE's student achievement to other countries, while both groups held negative perceptions about the extent to which LSAs provided information about the success of businesses on PE. Issues of teachers' practices of using provincial LSA results to determine and report student grades was raised by an interviewee, and again, highlighted parents' perception of teacher accountability. With respect to educational attainment, the ANOVA did not show significant difference in comparing mean scores for each of the five different educational levels.

In terms of participants' cultural affiliation, similar to the perceptions held by both parents and non-parents group, each cultural group did not believe that LSAs provided information about whether businesses were successful on the island. However, the large percentage (61.8%) of Newcomers who strongly believed that LSAs provided information about the quality of schools may be attributed to their school experiences and history of LSA in their homeland. This perception further illustrated in the scenario that more Newcomers (61%) and Aboriginal participants would move their child to a better performing school; a perception which was also shared by a large percentage of Grade 12 certificate and PhD holders.

Given the synthesis of the documented literature review and the research findings, the discussion explored possible factors influencing participants' knowledge about PE student performance on LSAs and their perceptions towards the utility of LSA as instrument of accountability. Although massive media reports and printed news about LSAs results in mathematics was regarded as the reason of the knowledge gained by participants. A more prominent presence of students' LSA performance in the media may be needed to inform the greater PE community.

Limitations

Limitations to this study focus on representation of participants, questionnaire items, and soliciting interview participants. Each of these limitations is discussed below.

Participants' responses were grouped into three variables: parental status, educational attainment, and cultural affiliations. There was sufficient representation by both parents and non-parents. However, when participants were grouped by educational attainment, there was more representation of participants with PhDs than the representation of PhD holders in

the PE population. At the same time, the number of participants with grade 12 certification or less were underrepresented in comparison to the greater PE population. Also, when participants were grouped by cultural affiliation, there was a greater percentage of Newcomers than the percentage in the PE population.

Upon reflecting on the questionnaire, there are many areas that could be changed. Firstly, more information should have been gathered on Newcomers to document their homeland and how long they have resided in PE. Also, it would have been worthy to know where (i.e., website, newspapers, etc.) that participants got their LSA information. Secondly, item 13 can be clarified to specify which LSAs that participants should focus on when responding to the sub-set of six items. A similar problem was presented in item 10 that asked participants to indicate how much they knew how PE scored on LSAs. This item followed an introduction to standardized tests listing the three levels of assessment (provincial, national, and international). Hence, this is a triple-barrelled question. Fixing each of these items would involve adding more items to explore each level of LSA separately which comes at a cost of increasing the length of the questionnaire, which may deter participants from completing it.

Another change focuses on the scenario presented in item 12 (Would you move your child to a better performing school if you could?). For this item, more information that would exclude obvious explanations preventing people from moving schools would more accurately capture whether participants' would move their child to a better performing school (if it was possible in every aspect).

In terms of the interview participants, including a teacher would have provided a better balance of perceptions. However, attempts were made to solicit a teacher to participate in

the study. In hindsight, some incentive may have been needed to encourage teachers to participate such as a $100 gift card. It is important to acknowledge that teachers are busy people and may need some incentive to encourage them to participate.

Implications and Recommendations

There are two primary implications for this research that focus on knowledge mobilization and gathering more research to extend the findings in this study. Two recommendations focus on accessibility of LSA at the provincial level and the need for an external testing body.

This thesis documented a number of areas or relationships that were not well understood by the public. The first relationship is between students' LSA performance and economic prosperity. This is probably the most poorly understood relationship in that participants in this study did not believe there is such a relationship. A strategy is needed to communicate to the public how or why LSAs are indicators of economic prosperity. Other relationships were between LSAs and school or teacher performance. Further, it is important that this information is made available in languages spoken by Newcomers to PE.

Findings in this thesis that call for more research include a focus on why the college/apprenticeship participants' knowledge about LSAs was less than their counterparts. In addition, there is a need to explore factors influencing why participants (parents in particular) chose not to move their child to a better performing school. It is possible that an improvement in the manner in which this question was posed would solicit a more accurate response, however, another round of responses are needed to verify and extend the findings for this item.

As shown in this thesis, PE's provincial LSA practices are not as transparent as other provinces. A recommendation for the Department of Education is to make all LSA reports available to the public. They also need to document student achievement according to the different levels of achievement so that the public can learn about the low number of students in the highest category and the large number of students performing in the lowest category, similar to PISA and PCAP (which should also be thoroughly reported on the provincial website). In addition, the Department of Education needs to clarify what the standards are so that the public knows what it means when, for example, 80% of the students met the standards. Lastly, the provincial government needs to consider creating an arms-length body (similar to the Educational and Quality Office in Ontario) to develop, administer, and analyze data for their LSA program to enhance transparency and conduct research using the data to inform the direction of schooling in the province.

References

Alberta Assessment Consortium. (2012). A new look at public assurance: Imagining the possibilities for Alberta students. Retrieved from http://www.aac.ab.ca/wp-content/uploads/2013/12/Public-Assurance-Doc_final_may31.pdf

Alberta Education. (2015a). General information bulletin: Diploma examinations program (2015-2016) school year. Retrieved from https://education.alberta.ca/media/1089201/2015-16-dip-gib-entire_20111118.pdf

Alberta Education. (2015b). Diploma exam accommodations. Retrieved from https://education.alberta.ca/media/1089194/08-dip-gib-2015-16-accommodations.pdf

Alberta Education. (2015c). School authority reporting. Retrieved from https://education.alberta.ca/school-authority-reporting/reporting-jurisdictions/

Alberta Teachers' Association. (2009). Real learning First: The teaching profession's view of student assessment, evaluation and accountability for the 21st century. *Issue in Education, 7*. Retrieved from http://www.teachers.ab.ca/SiteCollectionDocuments/ATA/Publications/Student-Assessment/PD-75-5%202009.pdf

Alibali, M. W., & Nathan, M. J. (2010). Conducting research in schools: A practical guide. *Journal of Cognition and Development, 11*(4), 397–407. Retrieved from http://2015.laschool4education.com/wp-content/uploads/2014/10/2010_AlibaliNathan_JCD_Tools.pdf

Armor, D., & Peiser, B. (1997). *Competition in education: A case study of interdistrict choice*. Boston: Pioneer Institute.

Bartholomew, D., Knott, M., & Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach (3rd ed.)*. West Sussex, UK: John Wiley & Sons.

Baum, S., Ma, J., & Payea, K. (2013). Education pays 2013: The benefits of higher education for individuals and society. Retrieved from https://trends.collegeboard.org/sites/default/files/education-pays-2013-full-report.pdf

Bhola, D. S., Buckendahl, C. W., & Impara, J. C., (2003). Aligning tests with states' content standards: Method and issues. *Educational Measurement Issues and Practice, 22*, 21–29.

Bolt, S. E., & Thurlow, M. L. (2004). Five of the most frequently allowed testing accommodations in State policy. *Remedial and Special Education, 25*(3), 141–152.

Boyce, C., & Neale, P. (2006). Pathfinder international tool series, conducting in-depth interviews: A guide for designing and conducting in-depth interview for evaluation input. *Pathfinder International Tool Series: Monitoring and Evaluation.* Retrieved from http://www2.pathfinder.org/site/DocServer/m_e_tool_series_indepth_interviews.pdf

Boyd, J. (2010). Early years evaluation sessions now taking place on PEI. Retrieved from http://www.boydjane.ca/2010/04/early-years-evalution-sessions-now-taking-place-on-pei/

Brookhart, S. M., & Durkin, D. T. (2003). Classroom assessment, student motivation, and achievement in high school social studies classes. *Applied Measurement In Education, 16*(1), 27–54.

Brophy, J. (2004). *Motivating students to learn*. Mahwah, NJ: Erlbaum.

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education, 11*, 301–318.

Brown, G. T. L., & Harris, L. R. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data. *Practical Assessment, Research and Evaluation, 15*(1). Retrieved from http://pareonline.net/pdf/v15n1.pdf

Burger, J. M., & Krueger, M. (2003). A balanced approach to high-stakes achievement testing: An analysis of the literature with policy implications. *International Electronic Journal For Learning, 7*(4). Retrieved from http://iejll.journalhosting.ucalgary.ca/index.php/ijll/article/view/413

CBC News. (2013, December 3). P.E.I. slips in international student testing. *CBC News*. Retrieved from http://www.cbc.ca/news/canada/prince-edward-island/p-e-i-slips-in-international-student-testing-1.2448740

CBC News. (2014, February 27). School choice greatest in Alberta, Fraser Institute says. *CBC News*. Retrieved from http://www.cbc.ca/news/canada/calgary/school-choice-greatest-in-alberta-fraser-institute-says-1.2553447

CBC News. (2014, March 4). From China to P.E.I.:The big picture. *CBC News*. Retrieved from http://www.cbc.ca/news/canada/prince-edward-island/from-china-to-p-e-i-the-big-picture-1.2557795

CBC News. (2015, November 5). P.E.I.'s English School Broad to be dissolved. *CBC News*. Retrieved from http://www.cbc.ca/news/canada/prince-edward-island/pei-education-administration-announcement-1.3305353

Child, D. (2006). *The essentials of factor analysis (3rd ed.).* New York, NY: Continuum International Publishing Group.

Childs, R. A., & Lawson, A. (2003). What do teacher candidates know about large-scale assessment? What should they know? *The Alberta Journal of Educational Research, XLIX* (4), 352–367. Retrieved from http://ajer.journalhosting.ucalgary.ca /index.php/ajer/article/viewFile/427/417

Christensen, L. B., & Johnson, R. B. (2013). *Educational research: Quantitative, qualitative and mixed approaches (5th ed.)*. Thousand Oaks, CA: Sage.

Cirtwill, C., Clifton, R., & D'orsay, J. (2002). *Testing and accountability: The keys to educational excellence in Atlantic Canada*. Halifax, NS: Atlantic Institute for Market Studies.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* New Jersey: Lawrence Erlbaum.

Columbia University. (2013). Pros and cons of standardized testing. Retrieved from http://worklife.columbia.edu/files_worklife/public/Pros_and_Cons_of_Standardized_Testing_1.pdf

Community and Health Services. (2014). A profile of baby-boomers and seniors in York region. Retrieved from http://www.york.ca/wps/wcm/connect/ yorkpublic/a3b167bb-3da1-4002-b294-abe88d71bdfd/ Profile_of_Boomers_and_Seniors_Jan_2014.pdf?MOD=AJPERES

Creswell, J. W. (2003). *Qualitative, quantitative, and mixed methods approaches (2nd ed.)*. Thousand Oaks, CA: Sage.

Cronenweth, S. (2012, January 4). High-stakes standardized testing in China. Retrieved from http://blog.socrato.com/high-stakes-standardized-testing-in-china/

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index

of consumer sentiment. *Public Opinion Quarterly 64*, 413–428.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1). Retrieved from https://depts.washington.edu/ctpmail/PDFs/LDH_1999.pdf

Davie, L. E., Hart, D. J., & Livingstone, D. W. (2001). *The 13th OISE survey: Public attitudes towards education in Ontario 2000.* Toronto: OISE Press.

DeCoster, J. (2006). *Testing Group Differences using T-tests, ANOVA, and Nonparametric Measures.* Retrieved from http://www.stathelp.com/notes.html

Demeris, H., Childs, R. A., & Jordan, A. (2007). The influence of students with special needs included in grade 3 classrooms on the large-scale achievement scores of students without special needs. *Canadian Journal of Education, 30*(3), 609–627.

Department of Education, Early Learning and Culture of Prince Edward Island. (2010). *Annual report 2009-2010*. Retrieved from http://www.gov.pe.ca/photos/original/eecd_AnReprt910.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2012). Early years evaluation. Retrieved from http://www.gov.pe.ca/eecd/index.php3?number=1031246&lang=E

Department of Education, Early Learning and Culture of Prince Edward Island. (2013). *The professional learning report*. Retrieved from http://www.gov.pe.ca/photos/original/ProfLearning_En.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2014). *Annual report 2013-2014*. Retrieved from http://www.gov.pe.ca/photos/original/eecd_AnRpt13-14.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2015a). Secondary mathematics assessment (SMA) Grade 11, information for parents. Retrieved from http://www.gov.pe.ca/photos/original/eelc_SMAPar16.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2015b). Provincial assessment program. Retrieved from http://www.gov.pe.ca/eecd/studentassessment

Department of Education, Early Learning and Culture of Prince Edward Island. (2015c). Elementary literacy assessment: Administration guide. Retrieved from http://www.gov.pe.ca/photos/original/eecd_ELAAdmin14.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2015d). School Act. Retrieved from http://www.gov.pe.ca/law/statutes/pdf/s-02_1.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2015e). Secondary literacy assessment (SLA) Grade 10. Retrieved from http://www.gov.pe.ca/eecd/index.php3?number=1053591&lang=E

Department of Education, Early Learning and Culture of Prince Edward Island. (2015f). Intermediate Literacy Assessment: Information for parents and guardians. Retrieved from http://www.gov.pe.ca/photos/original/eecd_ILA_Parent.pdf

Department of Education, Early Learning and Culture of Prince Edward Island. (2015g). Learning and early childhood development. Retrieved from http://www.gov.pe.ca/eecd/index.php3?number=1027651&lang=E

Department of Finance, Prince Edward Island Statistics Bureau. (2015). PEI unemployment rate. Retrieved from http://www.gov.pe.ca/photos/original/pt_lfs.pdf

DePascal, C. A. (2003). The ideal role of large-scale testing in a comprehensive assessment system. Retrieved from http://www.testpublishers.org /assets/documents/volume%205%20issue%201%20The%20ideal%20role.pdf

Driesler, S. D. (2001). Whiplash about backlash: The truth about public support for testing. *National Council on Measurement in Education Newsletter, 9*(3), 2–5.

Earl, L. M. (2010). Assessment and accountability in education: Improvement or surveillance? *Education Canada, 39*(3). Retrieved from http://www.cea-ace.ca/sites/default/files/EdCan-1999-v39-n3-Earl.pdf

Education Quality and Accountability Office (EQAO). (2014). *Annual report 2013-2014*. Retrieved from http://www.eqao.com/en/about_eqao/about_the_agency/ annual_reports/communication-docs/annual-report-2013-2014.pdf

Education Quality and Accountability Office (EQAO). (2015). *Ontario secondary school literacy test*. Retrieved from http://www.eqao.com/en/assessments/OSSLT

Elementary Teachers' Federation of Ontario (ETFO). (2015). *EQAO testing*. Retrieved from http://www.etfo.ca/issuesineducation/eqaotesting/Pages/default.aspx

Elliott, S. N., & Marquart, A. M. (2004). Extended time as a testing accommodation: Its effects and perceived consequences. *Council of Exceptional Children, 70*(3), 349–367.

Elmore, R. F. (2004). *School reform from the inside out: policy, practice, and performance*. Cambridge, MA: Harvard Education Press.

Etsey, Y. K. (1997, March). *Teachers' and school administrators' perspectives and use of standardized achievement tests: A review of published research*. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Falk, B., & Ort, S. (1998). Sitting down to score: Teacher learning through assessment.

Phi Delta Kappan, 80(1), 59–64.

Fey, P., & Smith, M. L. (2000). Validity and accountability of high-stakes testing. *Journal of Teacher Education, 51*(5), 334–344.

Freeman, K. A., & Williams, D. B. (2011, September). *Minority parents' perceptions of standardized student assessment*. Paper presented at the annual meeting of the European Educational Research Association on Educational Aspects of Rural Schooling: Possibilities, challenges, and changes. Budapest. Retrieved from http://www.eera-ecer.de/ecer-programmes/pdf/conference/5/contribution/20096/

Fushell, M. (2011, April). *Teachers' determination of final grades using large-scale assessments: Policy implications*. Paper presented at the annual meeting of the American Education Research Association. New Orleans.

Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C., Binkley, E., & Crouch, R. (2000). Using objective data sources to enhance teacher judgments about test accommodations. *Council of Exceptional Children, 67*(1), 67–81. doi:10.1177/001440290006700105

Galarneau, D. (1994). Baby boom women. *Perspectives, 6*(4). Retrieved from http://www.statcan.gc.ca/pub/75-001-x/1994004/1563-eng.pdf

Carlson, K. B. (2014, January 10). Math wars: The division over how to improve test scores. *The Global and Mail.* Retrieved from http://www.theglobeandmail.com/news/national/education/making-the-grade-the-provincial-push-for-better-math-scores/article16290351/?page=all

Garton, L., Haythornthwaite, C., & Wellman, B. (1999). Studying on-line social networks. In S. Jones (Ed.), *Doing Internet Research: Critical issues and methods for examining the net* (pp. 75–105). Thousand Oaks, CA: Sage.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies: A comparative analysis of six preconceptions about internet questionnaires. Retrived from http://www.simine.com/docs/Gosling_et_al_AP_2004.pdf

Gravetter, F. J., & Wallnau, L., B. (2004). *Statistics for the behavioral science*. Belmont, CA: Thomson/Wadsworth.

Greene, J. P., Peterson, P. E., & Du, J. (1998). School Choice in Milwaukee: A randomized experiment. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from School Choice* (pp. 335–356). Washington, DC: Brookings Institution Press.

Goodall, J., & Vorhaus, J. (2010). *Review of the best practice in parental engagement* (Report No. DFE-RR156). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182508/DFE-RR156.pdf

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies: A comparative analysis of six preconceptions about internet questionnaires. *American Psychological Association, 59*(2), 93–104.

Government of Prince Edward Island. (2013). *40th Annual Statistical Review*. Retrieved from http://www.gov.pe.ca/photos/original/pt_annualreview.pdf

Government of Prince Edward Island. (2014). *Highlights of Prince Edward Island. Prince Edward Island 2014 fall economic update*. Retrieved from http://www.gov.pe.ca/photos/original/leg_f14wsecoup.pdf

Government of Prince Edward Island. (2015a). *Prince Edward Island population report*.

Retrieved from http://www.gov.pe.ca/photos/original/pt_pop_rep.pdf

Government of Prince Edward Island. (2015b). *English language school board annual report 2014-2015*. Retrieved from http://www.gov.pe.ca/edu/elsb/files/2015/11/annual_report_2014_15.pdf

Government of Prince Edward Island. (2015c). New graduation requirements will better prepare students for future success. Retrieved from http://www.gov.pe.ca/index.php3/newsroom/index.php3?number=news&newsnumber=10103&dept=&lang=E

Cronenweth, S. (2012, January 4). High-stake standardized testing in China. Retrieved from http://blog.socrato.com/high-stakes-standardized-testing-in-china/

Harman, H. H. (1976). *Modern factor analysis (3rd ed. revised).* Chicago, IL: University of Chicago Press.

Hart, D. (2012). *Public attitudes towards education in Ontario 2012. The 18th OISE survey of educational issues*. Retrieved from https://www.oise.utoronto.ca/oise/UserFiles/File/OISE%20Survey/18th_OISE_Survey/OISE%20SURVEY%2018.pdf

Harris, D. N., & Sass, T. R. (2010). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, *95*(7-8), 798–812. Retrieved from http://www.caldercenter.org/sites/default/files/1001059_Teacher_Training.pdf

Hodnett, F. (2001). Evaluation versus assessment. Retrieved from http://pakeducationsite.com/education/articles/AssessmentandTesting/EvaluationVsAssessmet.doc

Holton, J. A. (2009). Qualitative tussles in undertaking a grounded theory study. *The Grounded Theory Review, 8*(3), 37–49.

Hopper, T. (2012, April 20). Is the PD day broken? Professional development days may do little to improve teaching. *National Post*. Retrieved from http://news.nationalpost.com/news/canada/is-the-pd-day-broken

Horne, P. (2008). Opening address. Prince Edward Island Teachers' Federation.

Huber, K. D, Rosenfeld, J. G., & Fiorello, C. A. (2001). The differential impact of inclusion and inclusive practices on high, average, and low achieving general education students. *Phycology in the Schools, 38*(6), 497–504. doi: 10.1002/pits. 1038

Huitt, W., Huitt, M., Monetti, D., & Hummel, J. (2009, October). *A systems-based synthesis of research related to improving students' academic performance*. Paper presented at the 3rd International City Break Conference sponsored by the Athens Institute for Education and Research (ATINER), Athens, Greece. Retrieved from http://www.edpsycinteractive.org/papers/improving-school-achievement.pdf

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*(6), 1258–1287.

Jones, K. (2004). A balanced school accountability model: An alternative to high-stakes testing. *Phi Delta Kappan 85*(8). 584–590. doi: 10.1177/003172170408500805 Retrieved from: http://pdk.sagepub.com/content/85/8/584.short

Jones, G., Jones, B., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impacts of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan 81*(3), 199–203.

Kao, L. S., & Green, C. E. (2008). Analysis of variance: Is there a difference in means and what does it mean? *Journal of Surgical Research, 144*(1), 150–170.

Kellaghan, T., & Madaus, G. F. (1992). Curriculum evaluation and assessment. In P. W. Jackson (Ed.), *Handbook of Research on Curriculum* (pp. 119–154). New York: Macmillan.

Kleitz, B., Weiher, G. R., Tedin, K., & Matland, R. (2000). Choice, charter schools, and household preferences. *Social Science Quarterly, 81*, 846–854.

Klinger, D. A., & Luce-Kapler R. (2007). Walking in their shoes: Students' perceptions of large-scale high-stake testing. *The Canadian of Program Evaluation, 22*(3), 29–52. Retrieved from http://www.evaluationcanada.ca/secure/22-3-029.pdf

Klinger, D. A., DeLuca, C, & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. *Canadian Journal of Educational Administration and Policy, 76*. Retrieved from https://www.umanitoba.ca/publications/cjeap/articles/klinger.html

Klinger, D. A., Maggi, S., & D'Angiulli, A. (2011). School accountability and assessment: Should we put the roof first? *The Educational Forum, 75*(2), 114–128.

Klinger. D. A., & Rogers, W. T. (2011). Teachers' perceptions of large-scale assessment programs within low-stakes accountability frameworks. *International Journal of Testing, 11*(2), 122–143.

Klinger, D. A., Volante, L., & DeLuca, C. (2012). Building teachers capacity within the evolving assessment culture in Canadian education. *Policy Futures in Education, 10*(4). Retrieved from http://pfe.sagepub.com/content/10/4/447.full.pdf

Klinger, D. A. (2015, April 22). As cited in Miller, T., Dropping assessment a mistake. *The*

Guardian. Retrieved from http://www.theguardian.pe.ca/Opinion/2015-04-22/article-4120870/Dropping-assessments-a-mistake/1

Klockars, A. J., & Sax, G. (1986). Post Hoc Comparisons: The Scheffe test. In A. J. Klockars & G. Sax (Eds.), *Multiple comparison: Issue 61* (pp. 58–61). CA: Sage

Koch, M. J. (2011, May). *Teachers' use of large-scale assessment as part of students' grades: Case studies of teacher practices in Ontario schools*. Paper presented at the Canadian Society for the Study of Education. UNB, Fredericton.

Kohn, A. (2001). Fighting the test: A practical guide to rescuing out schools. *Phi Delta Kappan, 82*(5), 349–357. Retrieved from http://www.alfiekohn.org/article/fighting-the-tests/

Koretz, D. M., & Barton, K. (2003). *Assessing students with disabilities: Issues and evidence* (Report No. 587). Retrieved from http://cse.ucla.edu/products/reports/TR587.pdf

Kraut, R., Olson, J., Mahzarin, B., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the internet. *American Psychological Association, 59*(2), 105–117.

Kurial, R. (2005). Excellence in education: A challenge for Prince Edward Island. Retrieved from http://www.gov.pe.ca/photos/original/task_force_edu.pdf

Landau, J. K., Vohs, J. R., & Romano, C. A. (1999). *Statewide assessment: Policy issues, questions, and strategies (parents engaged in education reform policy paper)*. Boston, MA: Federation for Children with Special Needs. Retrieved from www.fcsn.org/peer/ess/pdf/assessmentsw.pdf.

Legard, R., Keegan, J., & Ward, K. (2003). In-depth interviews, in qualitative research practice. In J. Ritchie & J. Lewis (Eds.) *A guide for social research students and researchers* (pp. 138–169), London: Sage.

Leithwood, K. (2006). Teacher working conditions that matter: Evidence for change. Retrieved from http://www.etfo.ca/Resources/ForTeachers/Documents/Teacher%20Working%20Conditions%20That%20Matter%20-%20Evidence%20for%20Change.pdf

Lenson, B. (2011, June 27). Standardized testing in China: What it says about American college prep. Retrieved from http://www.straighterline.com/blog/standardized-testing-in-china-what-it-says-about-american-college-prep/

Lewandowski, L. J., Lovett, B. J., Parolin, R., Gordon, M., & Codding, R. S. (2007). Extended time accommodations and the mathematics performance of students with and without ADHD. *Journal of Psychoeducational Assessment, 25*(1), 17–28.

Linn. R. L. (2000). Assessment and accountability. *Educational Researcher, 29*(2), 4–17.

Linn. R. L. (2003, July). *Accountability: Responsibility and seasonable expectations* (Report No. 601). Retrieved from http://www.cse.ucla.edu/products/reports/r601.pdf

Lipnevich, A. A., & Smith, J. K. (2008). *Response to assessment feedback: The effects of grades, praise, and source of information* (Report No. 08-30). Retrieved from https://www.ets.org/Media/Research/pdf/RR-08-30.pdf

Livingstone, D. W., Hart, D., & Davier, L. E. (2001). *Public attitudes towards education in Ontario 2000. The 13th OISE/UT survey*. Toronto: Ontario Institute for Studies in Education of the University of Toronto.

Manitoba Teachers' Society. (2010). Student Assessment: Standards tests have no academic value. Retrieved from http://www.mbteach.org/library/Archives/Issues/ISS-studentassessment.pdf

Matsuo, H., McIntyre, K. P., Tomazic, T., & Katz, B. (2004). The online survey: Its contributions and potential problems. *ASA Section on Survey Research Methods, 27*, 3998–4000.

McEwan, N. (1995). Accountability in education in Canada. *Canadian Journal of Education, 20*, 1–17.

McDonnell, M., Mclaughlin, M. J., & Morison, P. (1997). *Educating one and all* (Eds.). Washington, DC: National Academy Press.

McTighe, J., & Thomas, R. S. (2003). Backward design for forward action. *Educational Leadership, 60*(5), 52–55.

Mehrens, W. A. (1989). Preparing students to take standardized achievement test. *Practical Assessment, Research & Evaluation, 1*(11). Retrieved from http://pareonline.net/getvn.asp?v=1&n=11

Miller, T. (2013). Using large-scale assessment scores to determine student grades. *Canadian Journal of Education, 36*(3), 317–353. Retrieved from http://journals.sfu.ca/cje/index.php/cje-rce/article/viewFile/939/1609

Miller. T. (2015, April 22). Dropping assessment a mistake. *The Guardian*. Retrieved from http://www.theguardian.pe.ca/Opinion/2015-04-22/article-4120870/Dropping-assessments-a-mistake/1

Ministère de l'Éducation, du Loisir et du Sport. (2014). Exams: The final mark. Retrieved from http://www.education.gouv.qc.ca/en/parents-and-guardians/exams/

processing-of-results/the-final-mark/

Moore, D. L., & Tarnai, J. (2002). Evaluating nonresponse error in mail surveys. In: Groves, R. M., Dillman, D. A., Eltinge, J. L., & Little, R. J. A. (Eds.), *Survey Nonresponse* (pp. 197–211). New York: John Wiley & Sons

Moore, W. P., & Esselman, M. E. (1994). *Exploring the context of teacher efficacy: The role of achievement and climate*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Morison, P., McLaughlin, M. J., & McDonnell, L. M. (Eds.). (1997). *Educating One and All: Students with Disabilities and Standards-Based Reform*. Retrieved from http://files.eric.ed.gov/fulltext/ED409677.pdf

Mu, M., & Childs, R. (2005). What parents know and believe about large-scale assessments. *Canadian Journal of Educational Administration and Policy, 37*. Retrieved from https://www.umanitoba.ca/publications/cjeap/pdf_files/mu_childs.pdf

Muijs, R. D., & Rejnolds, D. (2002). Teachers' beliefs and behaviors: What really matters. *Journal of Classroom Interaction, 37*(2), 3−15.

Murry, T. S. (2003). Reflection of future international assessment. In D. S. Rychen, & L. H. Salgaanik (Eds.), *Key competencies for a successful life and a well-functioning society* (pp.135–160). Washington: Hogrefe & Hunber.

Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education, 25*, 262–279.

National Education Association. (2005). Culture abilities resilience effort: Strategies for closing the achievement gap. Retrieved from http://www.nea.org/assets/docs/mf_CAREbook0804.pdf

Newton, P., Tunison, S., & Viczko, M. (2010). The school principals' role large-scale assessment. *Canadian Journal of Educational Administration and Policy, 105*. Retrieved from https://www.umanitoba.ca/publications/cjeap/pdf_files/newton-tunison-viczko.pdf

Normand, R. (2008). School effectiveness or the horizon of the world as a laboratory. *British Journal of Sociology of Education, 29*(6), 665–676. doi: 10.1080/01425690802423346

Nunnally, J. C. (1978). *Psychometric theory (2nd ed.).* New York: McGraw-Hill.

Ontario Ministry of Education. (2003). The Ontario curriculum English. Retrieved from http://www.edu.gov.on.ca/eng/curriculum/secondary/english12curr.pdf

Ontario Ministry of Education. (2010). Growing success: Assessment, evaluation and reporting in Ontario schools. Retrieved from https://www.edu.gov.on.ca/eng/policyfunding/growSuccess.pdf

Ontario Ministry of Education. (2013). School effectiveness framework: A support for school improvement and student success. Retrieved from http://www.edu.gov.on.ca/eng/literacynumeracy/SEF2013.pdf

Ontario Ministry of Education. (2016). Using data: Transforming potential into practice. Retrieved from http://www.edu.gov.on.ca/eng/policyfunding/leadership/IdeasIntoActionBulletin5.pdf

Osburn. M. Z., Ritter, G., Stegman, C. S., & Suitt, L. D. (2004). Parents' perceptions of standardized testing: Is relationship and effect on student achievement. *Journal of Education Research & Policy Studies, 4*(1), 75–95. Retrieved from http://files.eric.ed.gov/fulltext/EJ848215.pdf

Parke, C. S. (2012). Making use of district and school data. *Practical Assessment, Research & Evaluation, 17*(10). Retrieved from http://pareonline.net/getvn.asp?v=17&n=10

PEI Home and School Federation. (2015, April 11). Leaders' debate on education (audio and questions). Retrieved from http://peihsf.ca/leaders-debate

Phelps, R. P. (1998). The demand for standardized student testing. *Educational Measurement: Issues and Practices, 17*(3), 5–23.

Phelps, R. P. (2005). Persistently positive: Forty years of public opinion on standardized testing. In R. P. Phelps (Ed.), *Defending standardized testing* (pp.1–22). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Popham, W. J. (2001). *The truth about testing: An educator's call to action*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

Prytula, M. P., Noonan, B., & Hellsten, L. M. (2013). Principals' perceptions of assessment leadership: A study of the assessment practices of school principals in Saskatchewan (Canada). *International Studies in Educational Administration (Commonwealth Council for Educational Administration & Management) (CCEAM), 40*(3), 57–74.

Rogers, W. T. (2014). Improving the utility of large-scale assessments in Canada. *Canadian Journal of Education, 37*(3). Retrieved from http://www.cje-rce.ca/index.php/cje-rce/article/view/1732/1703

Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education, 17*, 51–65.

Ross, J. A. (1998). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Advances in research on teaching* (pp. 49–74). Greenwich, CT: JAI Press.

Rossman, G. B., & Wilson. B. L. (1985). Number and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review 9*(5), 627–643.

Rowe, K. J. (2003). The importance of teacher quality as a key determinant of students' experiences and outcomes of schooling. Retrieved from http://www.nswteachers.nsw.edu.au/IgnitionSuite/uploads/docs/Rowe-The%20Importance%20of%20teacher%20Quality.pdf

Runyan, M. K. (1991). The effect of extra time on reading comprehension scores for university students with and without learning disabilities. *Journal of Learning Disabilities, 24*, 104–108.

Sacks, P. (2000). Predictable losers in testing schemes. *School Administrator, 57*(11), 6–9.

Samaroo, J., Dahya, N., & Alidina, S. (2013). Exploring the challenges of conducting respectful research: Seen and unforeseen factors within urban school research. *Canadian Journal of Education, 36*(3), 438–457. Retrieved from http://files.eric.ed.gov/fulltext/EJ1057944.pdf

Schleicher, A. (2010). The case for 21st century learning. Retrieved from http://www.oecd.org/general/thecasefor21st-centurylearning.htm

Sharp, M. N., York, J. L., & Knight, J. (1994). Effects of inclusion on the academic performance of classmates without disabilities. *Remedial and Special Education, 15* (5), 281–287.

Singer, E., van Hoewyk, J., & Maher, M. P. (2000). Experiments with incentives in telephone surveys. *Public Opinion Quarterly 64*, 171–188.

Simon, M., Ercikan, K., & Rousseau, M. (2013). *Improving large scale assessment in education: Theory, issues, and practice*. New York, NY: Taylor & Francis.

Skarupova, K. (2014). *Technical report: Computer-assisted and online data collection in general population surveys*. Lisbon: European Monitoring Center for Drugs and Drug Addiction. Retrieved from http://www.emcdda.europa.eu/publications/technical-reports/online-data-collection-gps

Statistics Canada. (2011). National occupational classification (NOC). Retrieved from http://www.statcan.gc.ca/pub/12-583-x/12-583-x2011001-eng.pdf

Statistics Canada. (2012). Online questionnaire: Reduced costs and increased quality. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/ref/overview-apercu/pop6-eng.cfm

Statistics Canada. (2013). Video: housing. Available from https://www12.statcan.gc.ca/nhs-enm/video/video-housing-logement-eng.cfm

Statistics Canada. (2015a). Focus on Geography series, 2011 census: Prince Edward Island. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-pr-eng.cfm?Lang=eng&GC=11

Statistics Canada. (2015b). *NHS Prince Edward Island, 2011*. Retrieved from https://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/prof/details/page.cfm?Lang=E&Geo1=HR&Code1=1100&Data=Count&SearchText=prince%20edward%20island&SearchType=Begins&SearchPR=11&A1=Labour&B1=All&Custom=&TABID=1

Statistics Canada. (2015c). *Labour force, employment and employment, levels and rates, by province*. Retrieved from http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/labor07a-eng.htm

Statistics Canada. (2015d). Median total income, by family type, by province and territory (All census families). Retrieved from http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/famil108a-eng.htm

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahway, NI: Erlbaum.

Stiggens, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan, 81*(3), 191–198.

Stiggins, R. J. (2001). The Unfulfilled Promise of Classroom Assessment. *Educational Measurement: Issues and Practice, 20*(3). Retrieved from http://www.edu.uwo.ca/source4allcourses/AQ/Intermediate_ABQ_New/downloads/stiggins.PDF

Strauss, V. (2014, April 23). How anxious are kids about taking standardized tests? This anxious. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2014/04/23/how-anxious-are-kids-about-taking-standardized-tests-this-anxious/

Strong, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (2008). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education, 20*(3), 165–184.

Sundgren, B. (2005). What is a public information system? Retrieved from http://www.diva-portal.org/smash/get/diva2:28973/FULLTEXT01.pdf

Taylor, A. R., & Tubianosa, T. (2001). *Student assessment in Canada: Improving the learning environment through effective evaluation*. Kelowna, BC: Society for the Advancement of Excellence in Education

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics (6th ed.).* Boston: Allyn and Bacon.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social & behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook on mixed methods in the behavioral and social sciences* (pp. 3–50). Thousand Oaks, CA: Sage.

The Conference Board of Canada. (2015a). Education Provincial rankings – How Canada performs. Retrieved from http://www.conferenceboard.ca/hcp/provincial/education.aspx

The Conferenced Board of Canada. (2015b). Education and learning. Retrieved from http://www.conferenceboard.ca/topics/education/default.aspx

The Council of Ministers of Education, Canada (CMEC). (1999). *School achievement indicators program 1998: Report on reading and writing assessment*. Retrieved from http://www.cmec.ca/Publications/Lists/Publications/Attachments/137/saip1998.en.pdf

The Council of Ministers of Education, Canada (CMEC). (2001). *Measuring up: The performance of Canada's youth in reading, mathematics and science OECD PISA Study—First Results for Canadians aged 15*. Retrieved from http://www.cmec.ca/docs/pisa2000/CanadaReport.EN.pdf

The Council of Ministers of Education, Canada (CMEC). (2004). *Measuring up: Canadian Results of the OECD PISA Study. The Performance of Canada's Youth in Mathematics, Reading, Science and Problem Solving 2003. First Findings for Canadians Aged 15.* Retrieved from http://www.cmec.ca/docs/pisa2003/pisa2003.en.pdf

The Council of Ministers of Education, Canada (CMEC). (2007*). Measuring up: Canadian Results of the OECD PISA Study. The Performance of Canada's Youth in Science, Reading and Mathematics 2006. First Results for Canadians Aged 15.* Retrieved from http://www.cmec.ca/docs/pisa2006/Pisa2006.en.pdf

The Council of Ministers of Education, Canada (CMEC). (2008). *PCAP-13 2007: Report on the assessment on 13-year-olds in reading, mathematics and science*. Retrieved from http://www.cmec.ca/Publications/Lists/Publications/Attachments/124/PCAP2007-Report.en.pdf

The Council of Ministers of Education, Canada (CMEC). (2010). *Measuring up: Canadian results of the OECD PISA Study. The performance of Canada's youth in reading, mathematics and science 2009: First results for Canadians aged 15.* Retrieved from http://www.cmec.ca/Publications/Lists/Publications/Attachments/254/PISA2009-can-report.pdf

The Council of Ministers of Education, Canada (CMEC). (2011). *PCAP-2010: Report on the Pan-Canadian assessment of mathematics, science and reading*. Retrieved from http://www.cmec.ca/Publications/Lists/Publications/Attachments/274/pcap2010.pdf

The Council of Ministers of Education, Canada (CMEC). (2013) *Measuring up: Canadian results of the OECD PISA results. The Performance of Canada's Youth in Reading,*

Mathematics and Science. 2012 First Results for Canadians Aged 15. Retrieved from http://cmec.ca/Publications/Lists/Publications/Attachments/318/PISA2012_CanadianReport_EN_Web.pdf

The Council of Ministers of Education, Canada (CMEC). (2014a). *PCAP-2013: Report on the Pan-Canadian assessment of science, reading and mathematics*. Retrieved from http://cmec.ca/Publications/Lists/Publications/Attachments/337/PCAP-2013-Public-Report-EN.pdf

The Council of Ministers of Education, Canada (CMEC). (2014b). *How good are Canadian 15-year-olds at solving problems? Further results from PISA 2012*. Retrieved from http://cmec.ca/Publications/Lists/Publications/Attachments/324/AMatters_No6_EN_Web.pdf

The Council of Ministers of Education, Canada (CMEC). (2015). *Pan-Canadian assessment program (PCAP)* overview. Retrieved from: http://www.cmec.ca/240/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-(PCAP)/Overview/index.html

The Forum of Educational Accountability. (2007). Assessment and accountability for improving schools and learning: Principals and recommendations for Federal law and State and local system. Retrieved from http://www.edaccountability.org/AssessmentFullReportJUNE07.pdf

The Guardian. (2015, April 12). Leaders response to questions of education. *The Guardian*. Retrieved from http://www.theguardian.pe.ca/Decision-15/2015-04-12/article-4108788/Leaders-respond-to-questions-of-education/1

The Organization of Economic and Co-operation Development (OECD). (2003). *Literacy*

skills for the world tomorrow: Further results from PISA 2000. Retrieved from http://www.oecd.org/edu/school/2960581.pdf

The Organization for Economic Co-operation and Development (OECD). (2007). *PISA 2006 Science Competencies for Tomorrow's World: Volume 1: Analysis*. doi: 10.1787/19962777. Retrieved from http://www.keepeek.com/Digital-Asset-Management/oecd/education/pisa-2006_9789264040014-en#page2

The Organization of Economic and Co-operation Development (OECD). (2013a). Education and skills online assessment. Retrieved from http://www.oecd.org/skills/ESonline-assessment/abouteducationskillsonline/EducationSkillsOnline%20Info%202%20(2).pdf

The Organization for Economic and Co-operation Development (OECD). (2013b). *PISA 2012 results: Excellence through equity – Giving every student the chance to succeed (Volume II)*. Paris: Author.

The Organization of Economic and Co-operation Development (OECD). (2014a). Education at a glance 2014: OECD indicators. Retrieved from http://www.oecd.org/edu/Education-at-a-Glance-2014.pdf

The Organization for Economic and Co-operation Development (OECD). (2014b). *PISA 2012 results in focus: What 15-year-olds know and what can they do with what they know.* Retrieved from http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Thompson, R. (2014, April 23). Too much test stress? Parents, experts discuss high-stake standardized test anxiety. *WJCT*. Retrieved from http://news.wjct.org/post/too-much-test-stress-parents-experts-discuss-high-stakes-standardized-test-anxiety

Toshalis, E., & Nakkula, M. J. (2012). Motivation, engagement, and student voice. Retrieved from http://www.studentsatthecenter.org/sites/scl.dl-dev.com/files/field_attach_file/Exec_Toshalis%26Nakkula_032312.pdf

van Barneveld, C. (2008). Using data to improve student achievement. *What works? Research into Practice*. Retrieved from https://www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/Using_Data.pdf

van Barneveld, C., & Brinson, K. (2011). *The rights and responsibilities of test takers when some parts of a large-scale test count toward class marks*. Paper presented at the American Education Research Association, New Orleans, LA.

van Barneveld, C., Pharand, S., Ruberto, L., & Haggarty, D. (2013). Motivation in largescale assessment. In Simon, M., Ercikan, K., & Rousseau, M. (Eds.). *Handbook on Large-scale Assessments : Theory and Practice* (pp.43–60). Thousand Oaks, CA: Sage Publications.

Vanourek, G., Bruno, V. M., Chester, E. F., Jr., & Louann A. B. (1998). Charter schools as seen by students, teachers, and parents. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from School Choice* (pp.187–211). Washington, DC: Brookings Institution Press.

Volante, L. (2004). Teaching to the test: What every educator and policy-makers should know. *Canadian Journal of Educational Administration and Policy, 35*. Retrieved from https://www.umanitoba.ca/publications/cjeap/articles/volante.html

Volante, L. (2006). Principles for effective classroom assessment. *Brock Education Journal, 15*(2), 134–147.

Volante, L. (2007). Educational quality and accountability in Ontario: Past, present, and future. *Canadian Journal of Educational Administration and Policy, 58*. Retrieved from https://umanitoba.ca/publications/cjeap/articles/volante_educational%20_quality.html

Volante, L. (2008). The impact of international comparison testing in Canada. *Journal of the International Society for Teacher Education, 12*(2), 17–20.

Volante, L. & Cherubini, L. (2010). Understanding the connections between large-scale assessment and school improvement planning. *Canadian Journal of Educational Administration and Policy, 115*. Retrieved from https://www.umanitoba.ca/publications/cjeap/pdf_files/volante-cherubini.pdf

Volante, L. (2014, July). *The revolution of international student assessment and educational policy development*. Paper presented at the International Conference on Learning, New York, NY.

von Davier, M., Gonzales, E., Kirsch, I., & Yamamoto, K. (Eds.) (2013). *The role of international large-scale assessments: Perspectives from technology, economy and educational research*. Dordrecht: Springer

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four States* (Research Monograph NO.18). Madison, WI: International Institution for Science Education.

Wellman, B. (1997). An electronic group is virtually a social network. In S. Kiesler (Ed.), *Culture of the internet* (pp. 179–205). Mahwah, NJ: Lawrence Erlbaum.

Wollack, J. A., & Fremer, J. J. (2013). *Handbook of test security*. New York, NY: Routledge.

Wright, T. (2012, May 22). PEI's Provincial Nominee Program a "serious failure," says Kenney. *The Guardian.* Retrieved online http://www.theguardian.pe.ca/News/Local/2012-05-22/article-2985604/P.E.I.s-PNP-a-serious-failure,-says-Kenney/1

Wright, T. (2013, August 31). PEI teachers' federation upset over money spent on standardized tests. *The Guardian*. Retrieved from http://www.theguardian.pe.ca/News/Local/2013-08-31/article-3370611/P.E.I.-Teachers%26rsquo%3B-Federation-upset-over-money-spent-on-standardized-tests/1

Wright, T. (2014a). Poor math scores in PEI have reached a crisis point. *The Guardian.* Retrieved from http://www.theguardian.pe.ca/News/Local/2014-01-20/article-3583962/Poor-math-scores-in-P.E.I.-have-reached-a-crisis-point%3A-UPEI-professor/1

Wright, T. (2014b). Poor math scores just don't add-up, opposition leaders says. *The Guardian.* Retrieved from http://www.theguardian.pe.ca/News/Local/2014-05-13/article-3724165/Poor-math-scores-among-P.E.I.-students-scrutinized-in-the-legislature/1

Zuriff, G. E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education, 3*, 99–117.

Zwaagstra, M. (2014, October, 16). Standardized tests benefit P.E.I. students. *The Guardian*. Retrieved from http://www.theguardian.pe.ca/Opinion/Letter-to-editor/2014-10-16/article-3905609/Standardized-tests-benefit-P.E.I.-students/

Appendix A

Questionnaire Sample

Background Information

1. We are trying to reach out to many people in the Prince Edward Island community. To help us understand the community you represent, please share details about your work and family.

Occupation _____

Employment Status _____

(full-time employed, part-time employed, self-employed, on leave, retired, unemployed, full-time student)

Parental Status (single parent/guardian, two parent/guardian, grandparent, n/a)

_____

2. Identify your gender.

☐ Female

☐ Male

3. Do you belong (e.g., do things with this cultural group) to one or more of the cultural groups listed below?

☐ French Community

☐ Aboriginal Community

☐ Newcomer to Prince Edward Island (new to PEI within the past 10 years from another province or country)

☐ Not associated with a cultural group

☐ Other

Please tell us the name of your cultural community. _____

4. Do you have a child in a public school (K to 12) on PEI now?

☐Yes ☐No

5. Do you have a grandchild in a public school (K to 12) on PEI now?

☐Yes ☐No

6. Identify your age range.

☐Pre baby boomer (born before 1945)

☐Born between 1945 and 1964

☐Post baby boomer (born after 1964)

7. Did you finish high school (grade 12)?

☐Yes ☐No

8. What is the highest level of education you have finished?

	Started but not able to finish	Working on it	Finished
Grade 12	☐	☐	☐
Apprenticeship	☐	☐	☐
College	☐	☐	☐
Undergraduate (e.g., BSc, BA)	☐	☐	☐
Graduate (e.g., Master)	☐	☐	☐
Graduate (e.g., PhD)	☐	☐	☐

9. Our goal is to get much feedback from people all over PEI. To know whether we met our goal, please tell us the name of the neighborhood of town where you live? (e.g., Sherwood, Tignish, Wilmot, etc.)

_____

Views on Standardized Test

All students in Prince Edward Island write a provincial test in Grade 3, 6 and 9. A grade 11 standardized test was just added. These tests are called **Common Assessments.** Grade 8 students in PEI write a Pan-Canadian Assessment Program (**PCAP**) test once every three years. This is a national test for students across Canada. Grade 10 students in PEI write the Program for International Student Assessment (**PISA**) test. This is an international test.

10. How much do you know about how PEI does on tests?

	Don't know	PEI is at or very near the top	Close to the top	In the middle	Close to the bottom	PEI is at or very near the bottom
(a) Reading						
(b) Math						
(c) Science						
(d) Problem-solving						

11. How important is the test in telling us about how well PEI students are doing?

	Don't know anything about it	1 Very important	2	3	4 Not important
(a) Provincial Test	☐	☐	☐	☐	☐
(b) National (PCAP) Test	☐	☐	☐	☐	☐
(c) International (PISA) Test	☐	☐	☐	☐	☐

12. If your child went to a school that had low scores on the provincial test, would you move them to a school that got better scores if you could?

☐Yes ☐No

13. How much do these tests tell us whether…

	1 Tell us a lot	2	3	4	5	6 Don't say anything
(a) the education system is working.	☐	☐	☐	☐	☐	☐
(b) students have the skills to be successful at college or university.	☐	☐	☐	☐	☐	☐
(c) children can solve problems in their own lives.	☐	☐	☐	☐	☐	☐
(d) students have the skills to take part in the world today.	☐	☐	☐	☐	☐	☐
(e) business are successful on the island.	☐	☐	☐	☐	☐	☐
(f) teachers are doing their job.	☐	☐	☐	☐	☐	☐

14. How much do you agree with each statement about the international PISA tests?

	1 Strongly Agree	2	3	4	5	6 Strongly Disagree	Don't know
(a) The tests tell us about children's success in school.	☐	☐	☐	☐	☐	☐	☐
(b) The tests are looked at by people thinking about moving to Canada.	☐	☐	☐	☐	☐	☐	☐
(c) The tests tell us about the quality of our schools.	☐	☐	☐	☐	☐	☐	☐
(d) The tests compare us to other countries.	☐	☐	☐	☐	☐	☐	☐
(e) The tests tell us about the quality of our teachers.	☐	☐	☐	☐	☐	☐	☐
(f) Businesses thinking about moving to Canada will	☐	☐	☐	☐	☐	☐	☐

look at the test scores.

15. How do PEI students compare with students in other provinces on the national and international tests?

	Below the provincial average	About the Same	Above the provincial average
(a) National (PCAP) Assessment	☐	☐	☐
(b) International (PISA) Assessment	☐	☐	☐

Interview Questions

General questions for all Participants:

1. Please tell me what you know about PEI student performance in the most recent provincial assessment/ national assessment/international assessment?
2. What do you think are the applications or uses of provincial assessment (common assessment), national assessment (PCAP) and international assessment (PISA) in PE?
3. One of the purposes of standardized tests is to measure the functionality of the education system. Do you think this is a good indicator of the education system? Why or why not?

Specific questions for provincial test developer

1. From examining the PISA questions and other provincial assessments used in other jurisdictions, we took a close look at the sample paper of Grade 9 Math test provided on the official website of Department of Education and Early Childhood Development (2013). How well does this test design align with teachers' classroom tests? How well does this test design align with the national assessment (PCAP) and the international assessment (PISA)?

2. Can you give me an example of the standards of the provincial assessment, for example, at each grade standard, what can students do? Below the standard, what can students do or not do (specifically)?

3. Can you give me an example of a parent report's about their child's provincial report?

4. With the trend of the "inclusive education" on provincial assessments, other provinces have initiated various ways meet the needs of students (different print types; audios, specially support), what are the solutions for PE students regarding to this aspect?

5. International standardized tests have begun to incorporate online testing. Similar initiatives are being piloted in provinces like Ontario. Describe any similar initiatives that PEI is currently doing or will be doing related to using technology to administer standardized tests.

6. According to EQAO, the largest third party that is responsible for the standardized assessments in Ontario, its annual cost to administer their provincial assessment is $17 per student, and it has a balanced assessment program providing 1) reliable data to guide effective teaching strategies; 2) information and evidence in planning school improvement; 3) identification of future success.

7. Do you know how much it cost to administer PEI's provincial (per student)? What is the financial model for provincial assessment?

8. According to Canadian statics, the high school dropout rate of PEI was 8.1% during the year 2009-2012 which is the same as the Canadian average. For post-secondary education, PEI had the highest rate of completion of post-secondary education (76%) and the lowest rate of dropout rate (9%), it also had the lowest rate of students pursuing further postsecondary education (8%, compared to 20% for Quebec, the province with the highest proportion) (Government of Canada, Minister of Industry, 2007). Based on the above statistics, do you think that the new Grade 11 Mathematics and Grade 10 literacy assessment will be good indicators of post-secondary attainment? For example, will a student who scores well in these two new tests most likely go to college or university?

9. From the provincial test data, we can see students made great improvement in their performance especially in math domain. However, from the national and international test results, PEI students are still at the low end of the scale. The PISA test focused on a) student's ability to apply what they learned in school to real-life situations; b) student learning motivations, self-beliefs and learning strategies; c) new skills help them to be adapt to 21st century. Given that PE students have not performed well, it seems that the curriculum or their education experience is not provide successful learning in these areas. What plans are being implemented to ensure that students receive the basic skills in a) real-life problem solving; b) learning strategies; c) new but necessary skills to take part in 21st century?

10. What evidence do you have that teachers are changing what they do in the classroom instructions based on provincial test results? For example, on Grade 9 mathematics provincial test in 2013-2014 school year, 71% of students were at or approaching the standards, while in 2013 Pan-Canadian assessment and 2012 PISA test, students did not score well, the scores were all below the Canadian average. What professional development did you plan for the teachers?

Specific Questions for high school principal/teacher

1. How do the standards in provincial assessment align with the standards used in classroom instructive assessments?

2. How do you/ the teachers in your school utilize the provincial assessment results for school improvement?

3. What are the effects of the provincial assessment in your school?

4. What are your most concerns about the new provincial assessment for Grade 10 and for Grade 11?

5. What do you think are the challenges that you/teachers are facing in practicing the provincial/ national/ international assessment?

6. What are the resources or supports that teachers and school administers need most to help improve student performance in provincial/ national/ international assessment?

7. Any other comments and suggestions that you would like to add in respect to the implication of provincial assessment at your school?

Specific Questions for representatives of businesses (e.g., bioscience)/Chamber of Commerce

1. With the new industrial investment and development in the growth sectors of aerospace, bioscience, information technology and renewable energy, what kind of skills should the future workforce be equipped with to meet the requirements in these newly developed industries?

2. What information (if any) does provincial, national, and international standardized tests provide businesses? What information should standardized tests provide for business and the larger community?

3. What are the factors influencing poor student achievement in PEI?

4. Explain the implications of poor PEI student achievement as it relates to businesses in general and more specifically, in your own business adventures.

5. What needs to be done to improve student achievement in PEI but also educate the children of PEI to be leaders the future?

Appendix B

Demographic Statistics Summary

We are trying to reach out to as many people in the Prince Edward Island community. To give help us understand the community you represent, please share details about your work and family.

D1. Indicate your occupation

66	(13.9%)	Business, finance and administrators
21	(4.4%)	Natural and applied sciences
29	(6.1%)	Health
235	(49.6%)	Education, law social, community workers
22	(4.6%)	Arts, culture, recreational services
48	(10.1%)	Sales and services
8	(1.7%)	Trade and transport
12	(2.5%)	Natural resources, agriculture
3	(0.6%)	Manufacturing and utilities
30	(6.3%)	Homemakers/stay at home mom

D2. Indicate your employment status

275	(55.9%)	Employed full time
53	(10.8%)	Employed part time
65	(13.2%)	Self employed
33	(6.7%)	Unemployed
30	(6.1%)	Retired
6	(1.2%)	Leave

11 (2.2%) Stay at home

19 (3.9%) Full time students

D3. Indicate your parental status

44 (9.8%) Single

33 (7.3%) Married (no kid)

52 (11.5%) Single Parent

289 (64.1%) Two parents

33 (7.3%) Grand parents

D4. Identify your gender.

376 (74.6%) Female

128 (25.4%) Male

D5. Do you belong (e.g. do things with this cultural group) to one or more of the cultural groups listed below?

37 (7.7%) French Community

18 (3.7%) Aboriginal Community

121 (25.1%) Newcomer to PEI (new to PEI within the past ten years from another province or country)

291 (60.4%) Not associated with a cultural group

15 (3.1%) Other

Please tell us the name of your cultural community----------

D6. Do you have a child in a public school (K to 12) on PEI now?

256 (50.6%) Yes 250 (49.4%) No

D7. Do you have a grandchild in a public school (K to 12) on PEI now?

29 (5.7%) Yes 478 (94.3%) No

D8. Identify your age range

14 (2.7%) Pre baby boomer (born before 1945)

144 (28.2%) Born between 1945 and 1964

353 (69.1%) Post baby boomer (born after 1964)

D9. Do you finish high school (grade 12)?

493 (99.4%) Yes 3 (0.6%) No

D10. What is the highest level of education you have finished?

	Started but unable to complete	Working on it	Completed
Below Grade 12	3 (0.6)	-	-
Grade 12	3 (1.5%)	(1%)	172 (98%)
College	13 (9%)	13 (9%)	114 (82%)
Undergraduate (e.g., BSc, BA)	18 (9%)	25 (12%)	167 (79%)
Graduate (e.g., Masters)	5 (3%)	16 (11%)	127 (86%)
Graduate (e.g., PhD)	7 (14%)	15 (29%)	29 (57%)

D11. Our goal is to get as much feedback from the people all over PEI.

To know whether we met our goal, please tell us the name of the neighbourhood or town where you live.

(e.g., Sherwood, Tignish, Wimot, etc.)

Appendix C

Descriptive Statistics Summary

Table C.1

Importance of the standardized tests in telling us how well PEI students are doing (Item 11).

	Don't Know	1 Very Important	2	3	4 Not Important	*M*	*SD*
Provincial	74 (14.5)	204 (40.1)	132 (25.9)	64 (12.6)	35 (6.9)	1.84	0.956
National	80 (15.7)	206 (40.5)	137 (26.9)	58 (11.4)	28 (5.5)	1.79	0.912
International	89 (17.5)	174 (34.3)	132 (26.0)	78 (15.4)	35 (6.9)	1.94	0.966

Table C.2

How much do these LSA tell us whether…(Item 13)

	Tell us a lot					… Don't say anything		
Item 13	1	2	3	4	5	6	*M*	*SD*
a. The education system is working.	173 (34.5)	102 (20.4)	107 (21.4)	44 (8.8)	34 (6.8)	41 (8.2)	2.57	1.58
b. Students have the skills to be successful at college or university.	125 (25.1)	99 (19.9)	114 (22.9)	67 (13.5)	44 (8.8)	49 (9.8)	2.91	1.60
c. Children can solve problems in their own lives.	82 (16.5)	64 (12.9)	135 (27.1)	75 (15.1)	62 (12.4)	80 (16.1)	3.42	1.65
d. Students have the skills to take part in the world today.	81 (16.3)	78 (15.7)	126 (25.4)	75 (15.1)	60 (12.1)	76 (15.3)	3.37	1.64
e. Businesses are successful on the island.	34 (8.4)	30 (7.4)	68 (16.9)	51 (12.7)	75 (18.6)	145 (36.0)	4.33	1.66
f. Teachers are doing their job.	109 (21.9)	86 (17.3)	118 (23.7)	67 (13.5)	53 (10.6)	65 (13.1)	3.13	1.67

Table C.3

How much do you agree with each statement about the international PISA tests (Item 14)?

Item 14	Strongly agree … 1	2	3	4	5	Strongly Disagree 6	Don't know	*M*	*SD*
a. The tests tell us about children's success in school.	96 (19.0)	72 (14.3)	112 (22.2)	73 (14.5)	51 (10.1)	44 (8.7)	56 (11.1)	3.10	1.586
b. The tests are looked by people thinking about moving to Canada.	105 (20.9)	53 (10.6)	80 (15.9)	34 (6.8)	29 (5.8)	34 (6.8)	167 (33.3)	2.79	1.655
c. The tests tell us about the quality of our schools.	104 (20.7)	86 (17.1)	97 (19.3)	60 (12.0)	49 (9.8)	51 (10.2)	55 (11)	3.04	1.651
d. The tests compare us to other countries.	189 (37.9)	111 (22.2)	66 (13.2)	29 (5.8)	17 (3.4)	12 (2.4)	75 (15.0)	2.08	1.299
e. The tests tell us about the quality of our teachers.	88 (17.7)	62 (12.5)	101 (20.4)	63 (12.7)	52 (10.5)	69 (13.9)	61 (12.3)	3.31	1.706
f. Businesses thinking about moving to Canada will look at the test scores.	63 (12.5)	55 (11.0)	69 (13.7)	48 (9.6)	48 (9.6)	44 (8.8)	175 (34.9)	3.29	1.676

Table C.4

Percentage of Participants' Feedback on Total Scale of LSA According to Parental Status

	Tell us a lot			.....	Don't say anything			
	1	2	3	4	5	6	*M*	*SD*
Item 13a								
Parents	93 (36.9)	54 (21.4)	51 (20.2)	21 (8.3)	17 (6.7)	16 (6.3)	2.46	1.526
Non-parents	77 (32.0)	47 (19.5)	54 (22.4)	21 (8.7)	17 (7.1)	25 (10.4)	2.71	1.643
Item 13b								
Parents	70 (27.8)	47 (18.7)	56 (22.2)	32 (12.7)	26 (10.3)	21 (8.3)	2.84	1.604
Non-parents	54 (22.7)	50 (21.0)	55 (23.1)	33 (13.9)	18 (7.6)	28 (11.8)	2.98	1.621
Item 13c								
Parents	44 (17.4)	32 (12.6)	69 (27.3)	40 (15.8)	31 (12.3)	37 (14.6)	3.37	1.632
Non-parents	37 (15.5)	31 (13.0)	62 (26.1)	35 (14.7)	31 (13.0)	42 (17.6)	3.50	1.666
Item 13d								
Parents	42 (15.2)	36 (14.3)	72 (28.7)	39 (15.5)	30 (12.0)	32 (12.7)	3.30	1.586
Non-parents	36 (15.2)	41 (17.3)	52 (21.9)	35 (14.8)	30 (12.7)	43 (18.1)	3.47	1.691
Item 13e								
Parents	14 (6.7)	17 (8.2)	39 (18.8)	29 (13.9)	43 (20.7)	66 (31.7)	4.29	1.589
Non-parents	18 (9.6)	13 (6.9)	28 (14.9)	22 (11.7)	29 (15.4)	78 (41.5)	4.41	1.717
Item 13f								
Parents	62 (24.6)	47 (18.7)	58 (23.0)	35 (13.9)	23 (9.1)	27 (10.7)	2.96	1.627
Non-parents	44 (18.5)	39 (16.4)	59 (24.8)	31 (13.0)	29 (12.2)	36 (15.1)	3.29	1.676
Item 14a								
Parents	56 (33.3)	23 (13.7)	37 (22.0)	21 (12.5)	14 (8.3)	17 (10.1)	3.04	1.602
Non-parents	46 (28.7)	29 (18.1)	41 (25.6)	13 (8.1)	15 (9.4)	16 (10.0)	3.17	1.559
Item 14b								
Parents	56 (33.3)	23 (13.7)	37 (22.0)	21 (12.5)	14 (8.3)	17 (10.1)	2.79	1.677
Non-parents	46 (28.7)	29 (18.1)	41 (25.6)	13 (8.1)	15 (9.4)	16 (10.0)	2.81	1.634
Item 14c								
Parents	59 (26.2)	43 (19.1)	48 (21.3)	32 (14.2)	26 (11.6)	17 (7.6)	2.88	1.588
Non-parents	42 (19.6)	42 (19.6)	47 (22.0)	28 (13.1)	22 (10.3)	33 (15.4)	3.21	1.697
Item 14d								
Parents	102 (49.3)	58 (28.0)	26 (12.6)	11 (5.3)	6 (2.9)	4 (1.9)	1.90	1.183
Non-parents	82 (39.2)	52 (24.9)	39 (18.7)	17 (8.1)	11(5.3)	8 (3.8)	2.27	1.388
Item 14e								
Parents	51 (23.3)	37 (16.9)	47 (21.5)	33 (15.1)	21 (9.6)	30 (13.7)	3.12	1.693
Non-parents	34 (16.3)	25 (12.0)	53 (25.5)	29 (13.9)	31 (14.9)	36 (17.3)	3.51	1.680
Item 14f								
Parents	35 (21.5)	23 (14.1)	34 (20.9)	23 (14.1)	24 (14.7)	24 (14.7)	3.31	1.726
Non-parents	27 (17.1)	31 (19.6)	33 (20.9)	25 (15.8)	24 (15.2)	18 (11.4)	3.27	1.618

Table C.5

Percentage of Participants' Feedback on Total Scale of LSA According to Educational Attainment

	Tell us a lot			...		Don't say anything		
	1	2	3	4	5	6	*M*	*SD*
Item 13a								
Grade 12	26 (52.0)	8 (16.0)	12 (24.0)	2 (4.0)	1 (2.0)	1 (2.0)	1.94	1.202
College/ Apprenticeship	43 (39.8)	22 (20.4)	22 (20.4)	8 (7.4)	4 (3.7)	9 (8.3)	2.40	1.558
Undergraduate	46 (32.6)	24 (17.0)	34 (24.1)	12 (8.5)	11 (7.8)	14 (9.9)	2.72	1.645
GraduateMaster	31 (24.0)	30 (23.3)	30 (23.3)	14 (10.9)	12 (9.3)	12 (9.3)	2.86	1.585
Graduate PhD	8 (28.6)	9 (32.1)	5 (17.9)	4 (14.3)	2 (7.1)	0 (0.0)	2.39	1.257
Item 13b								
Grade 12	17 (34.0)	10 (20.0)	13 (26.0)	5 (10.0)	2 (4.0)	3 (6.0)	2.48	1.460
College/ Apprenticeship	33 (30.6)	17 (15.7)	24 (22.2)	19 (17.6)	5 (4.6)	10 (9.3)	2.78	1.596
Undergraduate	34 (23.9)	28 (19.7)	33 (23.2)	16 (11.3)	13 (9.2)	18 (12.7)	3.00	1.672
GraduateMaster	22 (17.2)	31 (24.2)	33 (25.8)	17 (13.3)	13 (10.2)	12 (9.4)	3.03	1.526
Graduate PhD	8 (29.6)	6 (22.2)	3 (11.1)	4 (14.8)	5 (18.5)	1 (3.7)	2.81	1.642
Item 13c								
Grade 12	12 (24.0)	9 (18.0)	11 (22.0)	9 (18.0)	6 (12.0)	3 (6.0)	2.94	1.544
College/ Apprenticeship	22 (20.4)	12 (11.1)	27 (25.0)	13 (12.0)	13 (12.0)	21 (19.4)	3.43	1.768
Undergraduate	19 (13.4)	18 (12.7)	39 (27.5)	19 (13.4)	14 (9.9)	33 (23.2)	3.63	1.703
GraduatMaster	13 (10.2)	16 (12.6)	43 (33.9)	17 (13.4)	20 (15.7)	18 (14.2)	3.54	1.516
Graduate PhD	6 (21.4)	4 (14.3)	5 (17.9)	7 (25.0)	6 (21.4)	0 (0.0)	3.11	1.474
Item 13d								
Grade 12	12 (25.0)	10 (20.8)	9 (18.8)	6 (12.5)	8 (16.7)	3 (6.2)	2.94	1.616
College/ Apprenticeship	19 (17.6)	15 (13.9)	29 (26.9)	16 (14.8)	12 (11.1)	17 (15.7)	3.35	1.659
Undergraduate	20 (14.1)	21 (14.8)	39 (27.5)	20 (14.1)	13 (9.2)	29 (20.4)	3.51	1.679
GraduateMaster	14 (10.9)	21 (16.4)	38 (29.7)	18 (14.1)	19 (14.8)	18 (14.1)	3.48	1.547
Graduate PhD	6 (21.4)	6 (21.4)	3 (10.7)	6 (21.4)	4 (14.3)	3 (10.7)	3.18	1.701

(Table C.5 continued)

	Tell us a lot			...		Don't say anything		
	1	2	3	4	5	6	*M*	*SD*
Item 13e								
Grade 12	2 (8.3)	5 (20.8)	6 (25.0)	2 (8.3)	5 (20.8)	4 (16.7)	3.62	1.637
College/ Apprenticeship	10 (11.1)	3 (3.3)	19 (21.1)	11 (12.2)	14 (15.6)	13 (36.7)	4.28	1.709
Undergraduate	6 (5.1)	10 (8.5)	19 (16.2)	13 (11.1)	22 (18.8)	47 (40.2)	4.50	1.590
GraduateMaster	8 (6.8)	6 (5.1)	16 (13.6)	17 (14.4)	27 (22.9)	44 (37.3)	4.53	1.551
Graduate PhD	2 (9.1)	3 (13.6)	3 (13.6)	3 (13.6)	4 (18.2)	7 (31.8)	4.14	1.754
Item 13f								
Grade 12	22 (44.0)	8 (16.0)	9 (18.0)	3 (6.0)	5 (10.0)	3 (6.0)	2.40	1.616
College/ Apprenticeship	29 (27.1)	14 (13.1)	31 (29.0)	13 (12.1)	9 (8.4)	11 (10.3)	2.93	1.618
Undergraduate	20 (14.0)	25 (17.5)	37 (25.9)	19 (13.3)	15 (10.5)	27 (18.9)	3.45	1.673
GraduateMaster	19 (14.8)	26 (20.3)	30 (23.4)	18 (14.1)	19 (14.8)	16 (12.5)	3.31	1.606
Graduate PhD	6 (21.4)	5 (17.9)	6 (21.4)	6 (21.4)	3 (10.7)	2 (7.1)	3.04	1.551
Item 14a								
Grade 12	15 (34.9)	6 (14.0)	15 (34.9)	6 (14.0)	1 (2.3)	0 (0.0)	2.35	1.173
College/ Apprenticeship	19 (19.8)	20 (20.8)	21 (21.9)	19 (19.8)	8 (8.3)	9 (9.4)	3.04	1.549
Undergraduate	26 (20.5)	16 (12.6)	29 (22.8)	23 (18.1)	21 (16.5)	12 (9.4)	3.26	1.610
GraduateMaster	17 (14.5)	20 (17.1)	31 (26.5)	19 (16.2)	13 (11.1)	17 (14.5)	3.36	1.600
Graduate PhD	6 (23.1)	3 (11.5)	10 (38.5)	2 (7.7)	4 (15.4)	1 (3.8)	2.92	1.468
Item 14b								
Grade 12	19 (54.3)	6 (17.1)	4 (11.4)	3 (8.6)	0 (0.0)	3 (8.6)	2.09	1.560
College/ Apprenticeship	19 (26.8)	9 (12.7)	11 (15.5)	9 (12.7)	11 (15.5)	12 (16.9)	3.28	1.853
Undergraduate	31 (32.0)	14 (14.4)	23 (23.7)	13 (13.4)	7 (7.2)	9 (9.3)	2.77	1.630
GraduateMaster	19 (22.9)	18 (21.7)	26 (31.3)	5 (6.0)	7 (8.4)	8 (9.6)	2.84	1.550
Graduate PhD	6 (28.6)	3 (14.3)	7 (33.3)	1 (4.8)	4 (19.0)	0 (0.0)	2.71	1.454
Item 14c								
Grade 12	18 (40.9)	11 (25.0)	7 (15.9)	1 (2.3)	5 (11.4)	2 (4.5)	2.32	1.537
College/ Apprenticeship	24 (25.3)	16 (16.8)	24 (25.3)	14 (14.7)	9 (9.5)	8 (8.4)	2.92	1.575
Undergraduate	22 (17.5)	22 (17.5)	29 (23.0)	20 (15.9)	17 (13.5)	16 (12.7)	3.29	1.629
GraduateMaster	21 (17.9)	23 (19.7)	23 (19.7)	18 (15.4)	13 (11.1)	19 (16.2)	3.31	1.704
Graduate PhD	8 (29.6)	5 (18.5)	5 (18.5)	5 (18.5)	3 (11.1)	1 (3.7)	2.74	1.534
Item 14d								
Grade 12	24 (58.5)	8 (19.5)	5 (12.2)	3 (7.3)	1 (2.4)	0 (0.0)	1.76	1.090
College/ Apprenticeship	31 (38.3)	22 (27.2)	15 (18.5)	8 (9.9)	3 (3.7)	2 (2.5)	2.21	1.291
Undergraduate	48 (39.7)	37 (30.6)	21 (17.4)	9 (7.4)	4 (3.3)	2 (1.7)	2.09	1.197
GraduateMaster	50 (42.4)	30 (25.4)	17 (14.4)	8 (6.8)	6 (5.1)	7 (5.9)	2.25	1.485
Graduate PhD	16 (61.5)	7 (26.9)	2 (7.7)	0 (0.0)	1 (3.8)	0 (0.0)	1.58	0.945

(Table C.5 continued)

	Tell us a lot		...			Don't say anything		
	1	2	3	4	5	6	*M*	*SD*
Item 14e								
Grade 12	17 (39.5)	5 (11.6)	10 (23.3)	3 (7.0)	7 (16.3)	1 (2.3)	2.56	1.578
College/ Apprenticeship	19 (21.1)	10 (11.1)	24 (26.7)	16 (17.8)	10 (11.1)	11 (12.2)	3.23	1.629
Undergraduate	16 (12.8)	17 (13.6)	32 (25.6)	22 (17.6)	20 (16.0)	18 (14.4)	3.54	1.579
Graduate Master	20 (17.4)	15 (13.0)	26 (22.6)	13 (11.3)	11 (9.6)	30 (26.1)	3.61	1.824
Graduate PhD	5 (19.2)	6 (23.1)	4 (15.4)	5 (19.2)	2 (7.7)	4 (15.4)	3.19	1.721
Item 14f								
Grade 12	12 (35.3)	10 (29.4)	3 (8.8)	4 (11.8)	4 (11.8)	1 (2.9)	2.44	1.521
College/ Apprenticeship	10 (14.1)	9 (12.7)	11 (15.5)	12 (16.9)	14 (19.7)	15 (21.1)	3.79	1.723
Undergraduate	14 (15.4)	18 (19.8)	20 (22.0)	11 (12.1)	16 (17.6)	12 (13.2)	3.36	1.650
Graduate Master	14 (16.3)	13 (15.1)	19 (22.1)	15 (17.4)	10 (11.6)	15 (17.4)	3.45	1.685
Graduate PhD	5 (25.0)	4 (20.0)	7 (35.0)	3 (15.0)	1 (5.0)	0 (0.0)	2.55	1.191

Note. Participants with less than Grade 12 educational attainment were excluded due to the small sample.

Table C.6

Percentage of Participants' Feedback on Total Scale of LSA According to Cultural Affiliation

	Tell us a lot		...			Don't say anything		
	1	2	3	4	5	6	*M*	*SD*
Item 13a								
French	12 (37.5)	5 (15.6)	7 (21.9)	1 (3.1)	2 (6.2)	5 (15.6)	2.72	1.836
Aboriginal	5 (35.7)	1 (7.1)	1 (7.1)	3 (21.4)	2 (14.3)	2 (14.3)	3.14	1.956
Newcomers	41 (38.7)	21 (19.8)	22 (20.8)	8 (7.5)	4 (3.8)	10 (9.4)	2.46	1.599
Local	83 (31.4)	61 (23.1)	57 (21.6)	25 (9.5)	20 (7.6)	18 (6.9)	2.59	1.530
Other	6 (42.9)	2 (14.3)	4 (28.6)	1 (7.1)	1 (7.1)	0 (0.0)	2.21	1.311
Item 13b								
French	8 (25.0)	6 (18.8)	7 (21.9)	4 (12.5)	2 (6.2)	5 (15.6)	3.03	1.750
Aboriginal	5 (35.7)	0 (0.0)	1 (7.1)	6 (42.9)	0 (0.0)	2 (14.3)	3.14	1.834
Newcomers	32 (29.9)	20 (18.7)	20 (18.7)	13 (12.1)	11 (10.3)	11 (10.3)	2.85	1.687
Local	61 (23.3)	57 (21.8)	65 (24.8)	33 (12.6)	22 (8.4)	24 (9.2)	2.89	1.559
Other	5 (35.7)	2 (14.3)	4 (28.6)	2 (14.3)	1 (7.1)	0 (0.0)	2.43	1.342
Item 13c								
French	6 (18.8)	6 (18.8)	7 (21.9)	4 (12.5)	3 (9.4)	6 (18.8)	3.31	1.768
Aboriginal	1 (7.1)	2 (14.3)	2 (14.3)	2 (14.3)	2 (14.3)	5 (35.7)	4.21	1.762
Newcomers	17 (16.0)	17 (16.0)	24 (22.6)	16 (15.1)	16 (15.1)	16 (15.1)	3.42	1.662
Local	39 (14.8)	30 (11.4)	79 (30.0)	38 (14.4)	33 (12.5)	44 (16.7)	3.49	1.622
Other	6 (42.9)	1 (7.1)	4 (28.6)	1 (7.1)	1 (7.1)	1 (7.1)	2.50	1.653
Item 13d								
French	4 (12.5)	8 (25.0)	4 (12.5)	6 (18.8)	3 (9.4)	7 (21.9)	3.53	1.759
Aboriginal	0 (0.0)	5 (35.7)	1 (7.1)	3 (21.4)	1 (7.1)	4 (28.6)	3.86	1.703
Newcomers	16 (15.0)	21 (19.6)	23 (21.5)	16 (15.0)	15 (14.0)	16 (15.0)	3.38	1.652
Local	42 (16.0)	35 (13.4)	78 (29.8)	35 (13.4)	33 (12.6)	39 (14.9)	3.38	1.618
Other	6 (42.9)	1 (7.1)	3 (21.4)	2 (14.3)	1 (7.1)	1 (7.1)	2.57	1.697

(Table C.6 continued)

	Tell us a lot		…			Don't say anything		
	1	2	3	4	5	6	*M*	*SD*
Item 13e								
French	2 (6.5)	2 (6.5)	3 (9.7)	7 (22.6)	8 (25.8)	9 (29.0)	4.42	1.501
Aboriginal	0 (0.0)	1 (7.1)	4 (28.6)	2 (14.3)	3 (21.4)	4 (28.6)	4.36	1.393
Newcomers	6 (8.2)	9 (12.3)	6 (8.2)	12 (16.4)	18 (24.7)	22 (30.1)	4.27	1.652
Local	15 (6.8)	15 (6.8)	44 (19.9)	21 (9.5)	36 (16.3)	90 (40.7)	4.44	1.641
Other	2 (25.0)	0 (0.0)	0 (0.0)	2 (25.0)	2 (25.0)	2 (25.0)	4.00	2.000
Item 13f								
French	7 (21.9)	4 (12.5)	6 (18.8)	2 (6.2)	7 (21.9)	6 (18.8)	3.50	1.867
Aboriginal	4 (28.6)	1 (7.1)	3 (21.4)	1 (7.1)	3 (21.4)	2 (14.3)	3.29	1.899
Newcomers	29 (27.4)	22 (20.8)	20 (18.9)	13 (12.3)	10 (9.4)	12 (11.3)	2.90	1.684
Local	50 (18.9)	46 (17.4)	69 (26.1)	39 (14.8)	26 (9.8)	34 (12.9)	3.18	1.616
Other	2 (14.3)	4 (28.6)	6 (42.9)	0 (0.0)	1 (7.1)	1 (7.1)	2.79	1.369
Item 14a								
French	5 (17.2)	4 (13.8)	6 (20.7)	3 (10.3)	7 (24.1)	4 (13.8)	3.52	1.724
Aboriginal	0 (0.0)	2 (16.7)	4 (33.3)	3 (25.0)	2 (16.7)	1 (8.3)	3.67	1.231
Newcomers	26 (27.7)	17 (18.1)	23 (24.5)	13 (13.8)	7 (7.4)	8 (8.5)	2.81	1.575
Local	47 (19.5)	37 (15.4)	63 (26.1)	45 (18.7)	26 (10.8)	23 (9.5)	3.15	1.549
Other	4 (30.8)	3 (23.1)	3 (23.1)	0 (0.0)	2 (15.4)	1 (7.7)	2.69	1.702
Item 14b								
French	4 (15.4)	5 (19.2)	4 (15.4)	3 (11.5)	2 (7.7)	8 (30.8)	3.69	1.914
Aboriginal	2 (25.0)	1 (12.5)	1 (12.5)	2 (25.0)	2 (25.0)	0 (0.0)	3.13	1.642
Newcomers	23 (32.9)	16 (22.9)	20 (28.6)	4 (5.7)	2 (2.9)	5 (7.1)	2.44	1.441
Local	57 (32.0)	23 (12.9)	40 (22.5)	19 (10.7)	21 (11.8)	18 (10.1)	2.88	1.703
Other	4 (44.4)	3 (33.3)	2 (22.2)	0 (0.0)	0 (0.0)	0 (0.0)	1.78	0.833
Item 14c								
French	4 (13.3)	7 (23.3)	5 (16.7)	3 (10.0)	6 (20.0)	5 (16.7)	3.50	1.737
Aboriginal	2 (18.2)	1 (9.1)	3 (27.3)	2 (18.2)	2 (18.2)	1 (9.1)	3.36	1.629
Newcomers	32 (34.4)	24 (25.8)	13 (14.0)	11 (11.8)	6 (6.5)	7 (7.5)	2.53	1.578
Local	48 (19.9)	40 (16.6)	58 (24.1)	40 (16.6)	25 (10.4)	30 (12.4)	3.18	1.625
Other	4 (30.8)	4 (30.8)	1 (7.7)	0 (0.0)	2 (15.4)	2 (15.4)	2.85	1.951
Item 14d								
French	11 (36.7)	9 (30.0)	2 (6.7)	4 (13.3)	3 (10.0)	1 (3.3)	2.40	1.522
Aboriginal	36 (54.5)	2 (18.2)	0 (0.0)	3 (27.3)	0 (0.0)	0 (0.0)	2.00	1.342
Newcomers	48 (54.5)	15 (17.0)	15 (17.0)	6 (6.8)	1 (1.1)	3 (3.4)	1.93	1.285
Local	93 (41.3)	68 (30.2)	33 (14.7)	13 (5.8)	11 (4.9)	7 (3.1)	2.12	1.312
Other	5 (45.5)	4 (36.4)	2 (18.2)	0 (0.0)	0 (0.0)	0 (0.0)	1.73	0.786
Item 14e								
French	6 (21.4)	2 (7.1)	2 (7.1)	5 (17.9)	8 (28.6)	5 (17.9)	3.79	1.833
Aboriginal	2 (18.2)	1 (9.1)	2 (18.2)	4 (36.4)	0 (0.0)	2 (18.2)	3.45	1.695
Newcomers	25 (27.8)	14 (15.6)	23 (25.6)	11 (12.2)	8 (8.9)	9 (10.0)	2.89	1.631
Local	40 (16.9)	30 (12.7)	60 (25.4)	36 (15.3)	28 (11.9)	42 (17.8)	3.46	1.684
Other	3 (25.0)	3 (25.0)	3 (25.0)	0 (0.0)	1 (8.3)	2 (16.7)	2.92	1.832
Item 14f								
French	2 (7.4)	3 (11.1)	4 (14.8)	4 (14.8)	5 (18.5)	9 (33.3)	4.26	1.678
Aboriginal	1 (14.3)	2 (28.6)	0 (0.0)	2 (28.6)	2 (28.6)	0 (0.0)	3.29	1.604
Newcomers	17 (25.8)	15 (22.7)	11 (16.7)	14 (21.2)	4 (6.1)	5 (7.6)	2.82	1.548
Local	30 (16.9)	31 (17.4)	38 (21.3)	24 (13.5)	28 (15.7)	27 (15.7)	3.39	1.685
Other	4 (44.4)	2 (22.2)	3 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	1.89	0.928